# Linear Algebra

*and its applications*

FOURTH EDITION

David C. Lay

### SOLUTIONS TO PRACTICE PROBLEMS

**1.** Let $\mathbf{p}(t) = a_0 + a_1 t + a_2 t^2$ and compute

$$[T(\mathbf{p})]_{\mathcal{B}} = [T]_{\mathcal{B}}[\mathbf{p}]_{\mathcal{B}} = \begin{bmatrix} 3 & 4 & 0 \\ 0 & 5 & -1 \\ 1 & -2 & 7 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 3a_0 + 4a_1 \\ 5a_1 - a_2 \\ a_0 - 2a_1 + 7a_2 \end{bmatrix}$$

So $T(\mathbf{p}) = (3a_0 + 4a_1) + (5a_1 - a_2)t + (a_0 - 2a_1 + 7a_2)t^2$.

**2.** a. $A = (I)^{-1}AI$, so $A$ is similar to $A$.

b. By hypothesis, there exist invertible matrices $P$ and $Q$ with the property that $B = P^{-1}AP$ and $C = Q^{-1}BQ$. Substitute the formula for $B$ into the formula for $C$, and use a fact about the inverse of a product:

$$C = Q^{-1}BQ = Q^{-1}(P^{-1}AP)Q = (PQ)^{-1}A(PQ)$$

This equation has the proper form to show that $A$ is similar to $C$.

## 5.5 COMPLEX EIGENVALUES

Since the characteristic equation of an $n \times n$ matrix involves a polynomial of degree $n$, the equation always has exactly $n$ roots, counting multiplicities, *provided that possibly complex roots are included.* This section shows that if the characteristic equation of a real matrix $A$ has some complex roots, then these roots provide critical information about $A$. The key is to let $A$ act on the space $\mathbb{C}^n$ of $n$-tuples of complex numbers.[1]

Our interest in $\mathbb{C}^n$ does not arise from a desire to "generalize" the results of the earlier chapters, although that would in fact open up significant new applications of linear algebra.[2] Rather, this study of complex eigenvalues is essential in order to uncover "hidden" information about certain matrices with real entries that arise in a variety of real-life problems. Such problems include many real dynamical systems that involve periodic motion, vibration, or some type of rotation in space.

The matrix eigenvalue–eigenvector theory already developed for $\mathbb{R}^n$ applies equally well to $\mathbb{C}^n$. So a complex scalar $\lambda$ satisfies $\det(A - \lambda I) = 0$ if and only if there is a nonzero vector $\mathbf{x}$ in $\mathbb{C}^n$ such that $A\mathbf{x} = \lambda\mathbf{x}$. We call $\lambda$ a (**complex**) **eigenvalue** and $\mathbf{x}$ a (**complex**) **eigenvector** corresponding to $\lambda$.

**EXAMPLE 1** If $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$, then the linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ on $\mathbb{R}^2$ rotates the plane counterclockwise through a quarter-turn. The action of $A$ is periodic, since after four quarter-turns, a vector is back where it started. Obviously, no nonzero vector is mapped into a multiple of itself, so $A$ has no eigenvectors in $\mathbb{R}^2$ and hence no real eigenvalues. In fact, the characteristic equation of $A$ is

$$\lambda^2 + 1 = 0$$

[1] Refer to Appendix B for a brief discussion of complex numbers. Matrix algebra and concepts about real vector spaces carry over to the case with complex entries and scalars. In particular, $A(c\mathbf{x} + d\mathbf{y}) = cA\mathbf{x} + dA\mathbf{y}$, for $A$ an $m \times n$ matrix with complex entries, $\mathbf{x}, \mathbf{y}$ in $\mathbb{C}^n$, and $c, d$ in $\mathbb{C}$.

[2] A second course in linear algebra often discusses such topics. They are of particular importance in electrical engineering.

The only roots are complex: $\lambda = i$ and $\lambda = -i$. However, if we permit $A$ to act on $\mathbb{C}^2$, then

$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 1 \\ -i \end{bmatrix} = \begin{bmatrix} i \\ 1 \end{bmatrix} = i\begin{bmatrix} 1 \\ -i \end{bmatrix}$$

$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 1 \\ i \end{bmatrix} = \begin{bmatrix} -i \\ 1 \end{bmatrix} = -i\begin{bmatrix} 1 \\ i \end{bmatrix}$$

Thus $i$ and $-i$ are eigenvalues, with $\begin{bmatrix} 1 \\ -i \end{bmatrix}$ and $\begin{bmatrix} 1 \\ i \end{bmatrix}$ as corresponding eigenvectors. (A method for *finding* complex eigenvectors is discussed in Example 2.) ■

The main focus of this section will be on the matrix in the next example.

**EXAMPLE 2** Let $A = \begin{bmatrix} .5 & -.6 \\ .75 & 1.1 \end{bmatrix}$. Find the eigenvalues of $A$, and find a basis for each eigenspace.

**SOLUTION** The characteristic equation of $A$ is

$$\begin{aligned} 0 = \det\begin{bmatrix} .5-\lambda & -.6 \\ .75 & 1.1-\lambda \end{bmatrix} &= (.5-\lambda)(1.1-\lambda) - (-.6)(.75) \\ &= \lambda^2 - 1.6\lambda + 1 \end{aligned}$$

From the quadratic formula, $\lambda = \frac{1}{2}[1.6 \pm \sqrt{(-1.6)^2 - 4}] = .8 \pm .6i$. For the eigenvalue $\lambda = .8 - .6i$, construct

$$\begin{aligned} A - (.8 - .6i)I &= \begin{bmatrix} .5 & -.6 \\ .75 & 1.1 \end{bmatrix} - \begin{bmatrix} .8-.6i & 0 \\ 0 & .8-.6i \end{bmatrix} \\ &= \begin{bmatrix} -.3+.6i & -.6 \\ .75 & .3+.6i \end{bmatrix} \end{aligned} \tag{1}$$

Row reduction of the usual augmented matrix is quite unpleasant by hand because of the complex arithmetic. However, here is a nice observation that really simplifies matters: Since $.8 - .6i$ is an eigenvalue, the system

$$\begin{aligned} (-.3+.6i)x_1 - \quad\quad .6x_2 &= 0 \\ .75x_1 + (.3+.6i)x_2 &= 0 \end{aligned} \tag{2}$$

has a nontrivial solution (with $x_1$ and $x_2$ possibly complex numbers). Therefore, *both equations in (2) determine the same relationship between $x_1$ and $x_2$, and either equation can be used to express one variable in terms of the other.*[3]

The second equation in (2) leads to

$$\begin{aligned} .75x_1 &= (-.3 - .6i)x_2 \\ x_1 &= (-.4 - .8i)x_2 \end{aligned}$$

Choose $x_2 = 5$ to eliminate the decimals, and obtain $x_1 = -2 - 4i$. A basis for the eigenspace corresponding to $\lambda = .8 - .6i$ is

$$\mathbf{v}_1 = \begin{bmatrix} -2-4i \\ 5 \end{bmatrix}$$

[3] Another way to see this is to realize that the matrix in equation (1) is not invertible, so its rows are linearly dependent (as vectors in $\mathbb{C}^2$), and hence one row is a (complex) multiple of the other.

Analogous calculations for $\lambda = .8 + .6i$ produce the eigenvector

$$\mathbf{v}_2 = \begin{bmatrix} -2+4i \\ 5 \end{bmatrix}$$

As a check on the work, compute

$$A\mathbf{v}_2 = \begin{bmatrix} .5 & -.6 \\ .75 & 1.1 \end{bmatrix}\begin{bmatrix} -2+4i \\ 5 \end{bmatrix} = \begin{bmatrix} -4+2i \\ 4+3i \end{bmatrix} = (.8+.6i)\mathbf{v}_2 \qquad \blacksquare$$

Surprisingly, the matrix $A$ in Example 2 determines a transformation $\mathbf{x} \mapsto A\mathbf{x}$ that is essentially a rotation. This fact becomes evident when appropriate points are plotted.

**EXAMPLE 3** One way to see how multiplication by the matrix $A$ in Example 2 affects points is to plot an arbitrary initial point—say, $\mathbf{x}_0 = (2, 0)$—and then to plot successive images of this point under repeated multiplications by $A$. That is, plot

$$\begin{aligned}
\mathbf{x}_1 &= A\mathbf{x}_0 = \begin{bmatrix} .5 & -.6 \\ .75 & 1.1 \end{bmatrix}\begin{bmatrix} 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 1.0 \\ 1.5 \end{bmatrix} \\
\mathbf{x}_2 &= A\mathbf{x}_1 = \begin{bmatrix} .5 & -.6 \\ .75 & 1.1 \end{bmatrix}\begin{bmatrix} 1.0 \\ 1.5 \end{bmatrix} = \begin{bmatrix} -.4 \\ 2.4 \end{bmatrix} \\
\mathbf{x}_3 &= A\mathbf{x}_2, \ldots
\end{aligned}$$

Figure 1 shows $\mathbf{x}_0, \ldots, \mathbf{x}_8$ as larger dots. The smaller dots are the locations of $\mathbf{x}_9, \ldots, \mathbf{x}_{100}$. The sequence lies along an elliptical orbit. $\blacksquare$

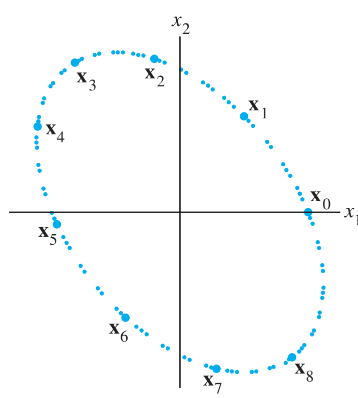



**FIGURE 1** Iterates of a point $\mathbf{x}_0$ under the action of a matrix with a complex eigenvalue.

Of course, Fig. 1 does not explain *why* the rotation occurs. The secret to the rotation is hidden in the real and imaginary parts of a complex eigenvector.

## Real and Imaginary Parts of Vectors

The complex conjugate of a complex vector $\mathbf{x}$ in $\mathbb{C}^n$ is the vector $\overline{\mathbf{x}}$ in $\mathbb{C}^n$ whose entries are the complex conjugates of the entries in $\mathbf{x}$. The **real** and **imaginary parts** of a complex vector $\mathbf{x}$ are the vectors $\text{Re}\,\mathbf{x}$ and $\text{Im}\,\mathbf{x}$ in $\mathbb{R}^n$ formed from the real and imaginary parts of the entries of $\mathbf{x}$.

**EXAMPLE 4** If $\mathbf{x} = \begin{bmatrix} 3-i \\ i \\ 2+5i \end{bmatrix} = \begin{bmatrix} 3 \\ 0 \\ 2 \end{bmatrix} + i \begin{bmatrix} -1 \\ 1 \\ 5 \end{bmatrix}$, then

$$\operatorname{Re}\mathbf{x} = \begin{bmatrix} 3 \\ 0 \\ 2 \end{bmatrix}, \quad \operatorname{Im}\mathbf{x} = \begin{bmatrix} -1 \\ 1 \\ 5 \end{bmatrix}, \quad \text{and} \quad \overline{\mathbf{x}} = \begin{bmatrix} 3 \\ 0 \\ 2 \end{bmatrix} - i \begin{bmatrix} -1 \\ 1 \\ 5 \end{bmatrix} = \begin{bmatrix} 3+i \\ -i \\ 2-5i \end{bmatrix} \quad \blacksquare$$

If $B$ is an $m \times n$ matrix with possibly complex entries, then $\overline{B}$ denotes the matrix whose entries are the complex conjugates of the entries in $B$. Properties of conjugates for complex numbers carry over to complex matrix algebra:

$$\overline{r\mathbf{x}} = \overline{r}\,\overline{\mathbf{x}}, \quad \overline{B\mathbf{x}} = \overline{B}\,\overline{\mathbf{x}}, \quad \overline{BC} = \overline{B}\,\overline{C}, \quad \text{and} \quad \overline{rB} = \overline{r}\,\overline{B}$$

## Eigenvalues and Eigenvectors of a Real Matrix That Acts on $\mathbb{C}^n$

Let $A$ be an $n \times n$ matrix whose entries are real. Then $\overline{A\mathbf{x}} = \overline{A}\,\overline{\mathbf{x}} = A\overline{\mathbf{x}}$. If $\lambda$ is an eigenvalue of $A$ and $\mathbf{x}$ is a corresponding eigenvector in $\mathbb{C}^n$, then

$$A\overline{\mathbf{x}} = \overline{A\mathbf{x}} = \overline{\lambda\mathbf{x}} = \overline{\lambda}\,\overline{\mathbf{x}}$$

Hence $\overline{\lambda}$ is also an eigenvalue of $A$, with $\overline{\mathbf{x}}$ a corresponding eigenvector. This shows that *when $A$ is real, its complex eigenvalues occur in conjugate pairs*. (Here and elsewhere, we use the term *complex eigenvalue* to refer to an eigenvalue $\lambda = a + bi$, with $b \neq 0$.)

**EXAMPLE 5** The eigenvalues of the real matrix in Example 2 are complex conjugates, namely, $.8 - .6i$ and $.8 + .6i$. The corresponding eigenvectors found in Example 2 are also conjugates:

$$\mathbf{v}_1 = \begin{bmatrix} -2-4i \\ 5 \end{bmatrix} \quad \text{and} \quad \mathbf{v}_2 = \begin{bmatrix} -2+4i \\ 5 \end{bmatrix} = \overline{\mathbf{v}}_1 \quad \blacksquare$$

The next example provides the basic "building block" for all real $2 \times 2$ matrices with complex eigenvalues.

**EXAMPLE 6** If $C = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$, where $a$ and $b$ are real and not both zero, then the eigenvalues of $C$ are $\lambda = a \pm bi$. (See the Practice Problem at the end of this section.) Also, if $r = |\lambda| = \sqrt{a^2 + b^2}$, then

$$C = r\begin{bmatrix} a/r & -b/r \\ b/r & a/r \end{bmatrix} = \begin{bmatrix} r & 0 \\ 0 & r \end{bmatrix} \begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix}$$

where $\varphi$ is the angle between the positive $x$-axis and the ray from $(0, 0)$ through $(a, b)$. See Fig. 2 and Appendix B. The angle $\varphi$ is called the *argument* of $\lambda = a + bi$. Thus the transformation $\mathbf{x} \mapsto C\mathbf{x}$ may be viewed as the composition of a rotation through the angle $\varphi$ and a scaling by $|\lambda|$ (see Fig. 3). $\blacksquare$

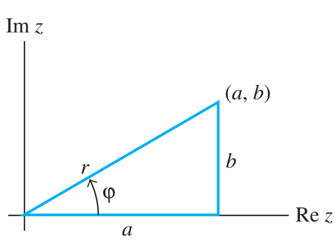


**FIGURE 2**

Finally, we are ready to uncover the rotation that is hidden within a real matrix having a complex eigenvalue.

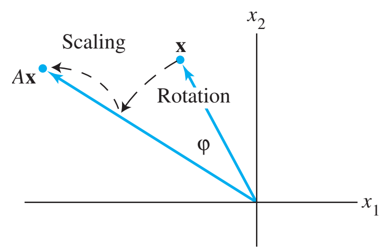



**FIGURE 3** A rotation followed by a scaling.

**EXAMPLE 7** Let $A = \begin{bmatrix} .5 & -.6 \\ .75 & 1.1 \end{bmatrix}$, $\lambda = .8 - .6i$, and $\mathbf{v}_1 = \begin{bmatrix} -2-4i \\ 5 \end{bmatrix}$, as in Example 2. Also, let $P$ be the $2 \times 2$ real matrix

$$P = \begin{bmatrix} \operatorname{Re}\mathbf{v}_1 & \operatorname{Im}\mathbf{v}_1 \end{bmatrix} = \begin{bmatrix} -2 & -4 \\ 5 & 0 \end{bmatrix}$$

and let

$$C = P^{-1}AP = \frac{1}{20}\begin{bmatrix} 0 & 4 \\ -5 & -2 \end{bmatrix}\begin{bmatrix} .5 & -.6 \\ .75 & 1.1 \end{bmatrix}\begin{bmatrix} -2 & -4 \\ 5 & 0 \end{bmatrix} = \begin{bmatrix} .8 & -.6 \\ .6 & 8 \end{bmatrix}$$

By Example 6, $C$ is a pure rotation because $|\lambda|^2 = (.8)^2 + (.6)^2 = 1$. From $C = P^{-1}AP$, we obtain

$$A = PCP^{-1} = P\begin{bmatrix} .8 & -.6 \\ .6 & 8 \end{bmatrix}P^{-1}$$

Here is the rotation "inside" $A$! The matrix $P$ provides a change of variable, say, $\mathbf{x} = P\mathbf{u}$. The action of $A$ amounts to a change of variable from $\mathbf{x}$ to $\mathbf{u}$, followed by a rotation, and then a return to the original variable. See Fig. 4. The rotation produces an ellipse, as in Fig. 1, instead of a circle, because the coordinate system determined by the columns of $P$ is not rectangular and does not have equal unit lengths on the two axes. ■

$$\begin{tikzcd} \mathbf{x} \arrow[r, "A"] \arrow[d, "P^{-1}", "\text{Change of variable}"'] & A\mathbf{x} \\ \mathbf{u} \arrow[r, "C"', "\text{Rotation}"] & C\mathbf{u} \arrow[u, "P"', "\text{Change of variable}"] \end{tikzcd}$$

**FIGURE 4** Rotation due to a complex eigenvalue.

The next theorem shows that the calculations in Example 7 can be carried out for any $2 \times 2$ real matrix $A$ having a complex eigenvalue $\lambda$. The proof uses the fact that if the entries in $A$ are real, then $A(\operatorname{Re}\mathbf{x}) = \operatorname{Re}A\mathbf{x}$ and $A(\operatorname{Im}\mathbf{x}) = \operatorname{Im}A\mathbf{x}$, and if $\mathbf{x}$ is an eigenvector for a complex eigenvalue, then $\operatorname{Re}\mathbf{x}$ and $\operatorname{Im}\mathbf{x}$ are linearly independent in $\mathbb{R}^2$. (See Exercises 25 and 26.) The details are omitted.

**THEOREM 9**

Let $A$ be a real $2 \times 2$ matrix with a complex eigenvalue $\lambda = a - bi$ $(b \neq 0)$ and an associated eigenvector $\mathbf{v}$ in $\mathbb{C}^2$. Then

$$A = PCP^{-1}, \quad \text{where} \quad P = \begin{bmatrix} \operatorname{Re}\mathbf{v} & \operatorname{Im}\mathbf{v} \end{bmatrix} \quad \text{and} \quad C = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$$

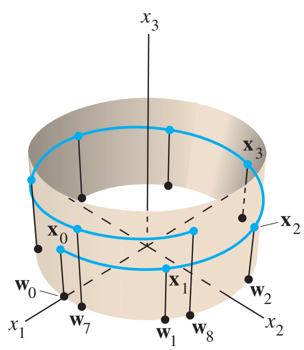



**FIGURE 5**

Iterates of two points under the action of a $3 \times 3$ matrix with a complex eigenvalue.

The phenomenon displayed in Example 7 persists in higher dimensions. For instance, if $A$ is a $3 \times 3$ matrix with a complex eigenvalue, then there is a plane in $\mathbb{R}^3$ on which $A$ acts as a rotation (possibly combined with scaling). Every vector in that plane is rotated into another point on the same plane. We say that the plane is **invariant** under $A$.

**EXAMPLE 8** The matrix $A = \begin{bmatrix} .8 & -.6 & 0 \\ .6 & .8 & 0 \\ 0 & 0 & 1.07 \end{bmatrix}$ has eigenvalues $.8 \pm .6i$ and 1.07. Any vector $\mathbf{w}_0$ in the $x_1x_2$-plane (with third coordinate 0) is rotated by $A$ into another point in the plane. Any vector $\mathbf{x}_0$ not in the plane has its $x_3$-coordinate multiplied by 1.07. The iterates of the points $\mathbf{w}_0 = (2, 0, 0)$ and $\mathbf{x}_0 = (2, 0, 1)$ under multiplication by $A$ are shown in Fig. 5. ■

## PRACTICE PROBLEM

Show that if $a$ and $b$ are real, then the eigenvalues of $A = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$ are $a \pm bi$, with corresponding eigenvectors $\begin{bmatrix} 1 \\ -i \end{bmatrix}$ and $\begin{bmatrix} 1 \\ i \end{bmatrix}$.

# 5.5 EXERCISES

Let each matrix in Exercises 1–6 act on $\mathbb{C}^2$. Find the eigenvalues and a basis for each eigenspace in $\mathbb{C}^2$.

**1.** $\begin{bmatrix} 1 & -2 \\ 1 & 3 \end{bmatrix}$

**2.** $\begin{bmatrix} 3 & -3 \\ 3 & 3 \end{bmatrix}$

**3.** $\begin{bmatrix} 5 & 1 \\ -8 & 1 \end{bmatrix}$

**4.** $\begin{bmatrix} 1 & -2 \\ 1 & 3 \end{bmatrix}$

**5.** $\begin{bmatrix} 3 & 1 \\ -2 & 5 \end{bmatrix}$

**6.** $\begin{bmatrix} 7 & -5 \\ 1 & 3 \end{bmatrix}$

In Exercises 7–12, use Example 6 to list the eigenvalues of $A$. In each case, the transformation $\mathbf{x} \mapsto A\mathbf{x}$ is the composition of a rotation and a scaling. Give the angle $\varphi$ of the rotation, where $-\pi < \varphi \le \pi$, and give the scale factor $r$.

**7.** $\begin{bmatrix} \sqrt{3} & -1 \\ 1 & \sqrt{3} \end{bmatrix}$

**8.** $\begin{bmatrix} 3 & 3\sqrt{3} \\ -3\sqrt{3} & 3 \end{bmatrix}$

**9.** $\begin{bmatrix} 0 & 2 \\ -2 & 0 \end{bmatrix}$

**10.** $\begin{bmatrix} 0 & .5 \\ -.5 & 0 \end{bmatrix}$

**11.** $\begin{bmatrix} -\sqrt{3} & 1 \\ -1 & -\sqrt{3} \end{bmatrix}$

**12.** $\begin{bmatrix} 3 & -\sqrt{3} \\ \sqrt{3} & 3 \end{bmatrix}$

In Exercises 13–20, find an invertible matrix $P$ and a matrix $C$ of the form $\begin{bmatrix} a & -b \\ b & a \end{bmatrix}$ such that the given matrix has the form $A = PCP^{-1}$.

**13.** $\begin{bmatrix} 1 & -2 \\ 1 & 3 \end{bmatrix}$

**14.** $\begin{bmatrix} 3 & -3 \\ 1 & 1 \end{bmatrix}$

**15.** $\begin{bmatrix} 0 & 5 \\ -2 & 2 \end{bmatrix}$

**16.** $\begin{bmatrix} 4 & -2 \\ 1 & 6 \end{bmatrix}$

**17.** $\begin{bmatrix} -11 & -4 \\ 20 & 5 \end{bmatrix}$

**18.** $\begin{bmatrix} 3 & -5 \\ 2 & 5 \end{bmatrix}$

**19.** $\begin{bmatrix} 1.52 & -.7 \\ .56 & .4 \end{bmatrix}$

**20.** $\begin{bmatrix} -3 & -8 \\ 4 & 5 \end{bmatrix}$

**21.** In Example 2, solve the first equation in (2) for $x_2$ in terms of $x_1$, and from that produce the eigenvector $\mathbf{y} = \begin{bmatrix} 2 \\ -1 + 2i \end{bmatrix}$ for the matrix $A$. Show that this $\mathbf{y}$ is a (complex) multiple of the vector $\mathbf{v}_1$ used in Example 2.

**22.** Let $A$ be a complex (or real) $n \times n$ matrix, and let $\mathbf{x}$ in $\mathbb{C}^n$ be an eigenvector corresponding to an eigenvalue $\lambda$ in $\mathbb{C}$. Show that for each nonzero complex scalar $\mu$, the vector $\mu\mathbf{x}$ is an eigenvector of $A$.

Chapter 7 will focus on matrices $A$ with the property that $A^T = A$. Exercises 23 and 24 show that every eigenvalue of such a matrix is necessarily real.

**23.** Let $A$ be an $n \times n$ real matrix with the property that $A^T = A$, let $\mathbf{x}$ be any vector in $\mathbb{C}^n$, and let $q = \bar{\mathbf{x}}^T A\mathbf{x}$. The equalities below show that $q$ is a real number by verifying that $\bar{q} = q$. Give a reason for each step.

$$\bar{q} = \overline{\bar{\mathbf{x}}^T A\mathbf{x}} \underset{\text{(a)}}{=} \mathbf{x}^T \overline{A\mathbf{x}} \underset{\text{(b)}}{=} \mathbf{x}^T A\bar{\mathbf{x}} \underset{\text{(c)}}{=} (\mathbf{x}^T A\bar{\mathbf{x}})^T \underset{\text{(d)}}{=} \bar{\mathbf{x}}^T A^T\mathbf{x} \underset{\text{(e)}}{=} q$$

**24.** Let $A$ be an $n \times n$ real matrix with the property that $A^T = A$. Show that if $A\mathbf{x} = \lambda\mathbf{x}$ for some nonzero vector $\mathbf{x}$ in $\mathbb{C}^n$, then, in fact, $\lambda$ is real and the real part of $\mathbf{x}$ is an eigenvector of $A$. [*Hint:* Compute $\overline{\mathbf{x}}^T A\mathbf{x}$, and use Exercise 23. Also, examine the real and imaginary parts of $A\mathbf{x}$.]

**25.** Let $A$ be a real $n \times n$ matrix, and let $\mathbf{x}$ be a vector in $\mathbb{C}^n$. Show that $\text{Re}(A\mathbf{x}) = A(\text{Re}\,\mathbf{x})$ and $\text{Im}(A\mathbf{x}) = A(\text{Im}\,\mathbf{x})$.

**26.** Let $A$ be a real $2 \times 2$ matrix with a complex eigenvalue $\lambda = a - bi$ $(b \neq 0)$ and an associated eigenvector $\mathbf{v}$ in $\mathbb{C}^2$.

a. Show that $A(\text{Re}\,\mathbf{v}) = a\,\text{Re}\,\mathbf{v} + b\,\text{Im}\,\mathbf{v}$ and $A(\text{Im}\,\mathbf{v}) = -b\,\text{Re}\,\mathbf{v} + a\,\text{Im}\,\mathbf{v}$. [*Hint:* Write $\mathbf{v} = \text{Re}\,\mathbf{v} + i\,\text{Im}\,\mathbf{v}$, and compute $A\mathbf{v}$.]

b. Verify that if $P$ and $C$ are given as in Theorem 9, then $AP = PC$.

[**M**] In Exercises 27 and 28, find a factorization of the given matrix $A$ in the form $A = PCP^{-1}$, where $C$ is a block-diagonal matrix with $2 \times 2$ blocks of the form shown in Example 6. (For each conjugate pair of eigenvalues, use the real and imaginary parts of one eigenvector in $\mathbb{C}^4$ to create two columns of $P$.)

**27.** $A = \begin{bmatrix} 26 & 33 & 23 & 20 \\ -6 & -8 & -1 & -13 \\ -14 & -19 & -16 & 3 \\ -20 & -20 & -20 & -14 \end{bmatrix}$

**28.** $A = \begin{bmatrix} 7 & 11 & 20 & 17 \\ -20 & -40 & -86 & -74 \\ 0 & -5 & -10 & -10 \\ 10 & 28 & 60 & 53 \end{bmatrix}$

### SOLUTION TO PRACTICE PROBLEM

Remember that it is easy to test whether a vector is an eigenvector. There is no need to examine the characteristic equation. Compute

$$A\mathbf{x} = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}\begin{bmatrix} 1 \\ -i \end{bmatrix} = \begin{bmatrix} a+bi \\ b-ai \end{bmatrix} = (a+bi)\begin{bmatrix} 1 \\ -i \end{bmatrix}$$

Thus $\begin{bmatrix} 1 \\ -i \end{bmatrix}$ is an eigenvector corresponding to $\lambda = a + bi$. From the discussion in this section, $\begin{bmatrix} 1 \\ i \end{bmatrix}$ must be an eigenvector corresponding to $\overline{\lambda} = a - bi$.

## 5.6 DISCRETE DYNAMICAL SYSTEMS

Eigenvalues and eigenvectors provide the key to understanding the long-term behavior, or *evolution*, of a dynamical system described by a difference equation $\mathbf{x}_{k+1} = A\mathbf{x}_k$. Such an equation was used to model population movement in Section 1.10, various Markov chains in Section 4.9, and the spotted owl population in the introductory example for this chapter. The vectors $\mathbf{x}_k$ give information about the system as time (denoted by $k$) passes. In the spotted owl example, for instance, $\mathbf{x}_k$ listed the numbers of owls in three age classes at time $k$.

The applications in this section focus on ecological problems because they are easier to state and explain than, say, problems in physics or engineering. However, dynamical systems arise in many scientific fields. For instance, standard undergraduate courses in control systems discuss several aspects of dynamical systems. The modern *state-space* design method in such courses relies heavily on matrix algebra.[1] The *steady-state response* of a control system is the engineering equivalent of what we call here the "long-term behavior" of the dynamical system $\mathbf{x}_{k+1} = A\mathbf{x}_k$.

[1] See G. F. Franklin, J. D. Powell, and A. Emami-Naeimi, *Feedback Control of Dynamic Systems*, 5th ed. (Upper Saddle River, NJ: Prentice-Hall, 2006). This undergraduate text has a nice introduction to dynamic models (Chapter 2). State-space design is covered in Chapters 7 and 8.

# 7.1 DIAGONALIZATION OF SYMMETRIC MATRICES

A **symmetric** matrix is a matrix $A$ such that $A^T = A$. Such a matrix is necessarily square. Its main diagonal entries are arbitrary, but its other entries occur in pairs—on opposite sides of the main diagonal.

**EXAMPLE 1** Of the following matrices, only the first three are symmetric:

$$\text{Symmetric:} \quad \begin{bmatrix} 1 & 0 \\ 0 & -3 \end{bmatrix}, \quad \begin{bmatrix} 0 & -1 & 0 \\ -1 & 5 & 8 \\ 0 & 8 & -7 \end{bmatrix}, \quad \begin{bmatrix} a & b & c \\ b & d & e \\ c & e & f \end{bmatrix}$$

$$\text{Nonsymmetric:} \quad \begin{bmatrix} 1 & -3 \\ 3 & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & -4 & 0 \\ -6 & 1 & -4 \\ 0 & -6 & 1 \end{bmatrix}, \quad \begin{bmatrix} 5 & 4 & 3 & 2 \\ 4 & 3 & 2 & 1 \\ 3 & 2 & 1 & 0 \end{bmatrix}$$

■

To begin the study of symmetric matrices, it is helpful to review the diagonalization process of Section 5.3.

**EXAMPLE 2** If possible, diagonalize the matrix $A = \begin{bmatrix} 6 & -2 & -1 \\ -2 & 6 & -1 \\ -1 & -1 & 5 \end{bmatrix}$.

**SOLUTION** The characteristic equation of $A$ is

$$0 = -\lambda^3 + 17\lambda^2 - 90\lambda + 144 = -(\lambda - 8)(\lambda - 6)(\lambda - 3)$$

Standard calculations produce a basis for each eigenspace:

$$\lambda = 8: \ \mathbf{v}_1 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}; \qquad \lambda = 6: \ \mathbf{v}_2 = \begin{bmatrix} -1 \\ -1 \\ 2 \end{bmatrix}; \qquad \lambda = 3: \ \mathbf{v}_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

These three vectors form a basis for $\mathbb{R}^3$. In fact, it is easy to check that $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is an *orthogonal* basis for $\mathbb{R}^3$. Experience from Chapter 6 suggests that an *orthonormal* basis might be useful for calculations, so here are the normalized (unit) eigenvectors.

$$\mathbf{u}_1 = \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix}, \quad \mathbf{u}_2 = \begin{bmatrix} -1/\sqrt{6} \\ -1/\sqrt{6} \\ 2/\sqrt{6} \end{bmatrix}, \quad \mathbf{u}_3 = \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}$$

Let

$$P = \begin{bmatrix} -1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \\ 1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \\ 0 & 2/\sqrt{6} & 1/\sqrt{3} \end{bmatrix}, \quad D = \begin{bmatrix} 8 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

Then $A = PDP^{-1}$, as usual. But this time, since $P$ is square and has orthonormal columns, $P$ is an *orthogonal* matrix, and $P^{-1}$ is simply $P^T$. (See Section 6.2.) ■

Theorem 1 explains why the eigenvectors in Example 2 are orthogonal—they correspond to distinct eigenvalues.

**THEOREM 1**

If $A$ is symmetric, then any two eigenvectors from different eigenspaces are orthogonal.

**PROOF** Let $\mathbf{v}_1$ and $\mathbf{v}_2$ be eigenvectors that correspond to distinct eigenvalues, say, $\lambda_1$ and $\lambda_2$. To show that $\mathbf{v}_1 \cdot \mathbf{v}_2 = 0$, compute

$$\begin{aligned}
\lambda_1 \mathbf{v}_1 \cdot \mathbf{v}_2 &= (\lambda_1 \mathbf{v}_1)^T \mathbf{v}_2 = (A\mathbf{v}_1)^T \mathbf{v}_2 && \text{Since } \mathbf{v}_1 \text{ is an eigenvector} \\
&= (\mathbf{v}_1^T A^T)\mathbf{v}_2 = \mathbf{v}_1^T (A\mathbf{v}_2) && \text{Since } A^T = A \\
&= \mathbf{v}_1^T (\lambda_2 \mathbf{v}_2) && \text{Since } \mathbf{v}_2 \text{ is an eigenvector} \\
&= \lambda_2 \mathbf{v}_1^T \mathbf{v}_2 = \lambda_2 \mathbf{v}_1 \cdot \mathbf{v}_2
\end{aligned}$$

Hence $(\lambda_1 - \lambda_2)\mathbf{v}_1 \cdot \mathbf{v}_2 = 0$. But $\lambda_1 - \lambda_2 \neq 0$, so $\mathbf{v}_1 \cdot \mathbf{v}_2 = 0$. ■

The special type of diagonalization in Example 2 is crucial for the theory of symmetric matrices. An $n \times n$ matrix $A$ is said to be **orthogonally diagonalizable** if there are an orthogonal matrix $P$ (with $P^{-1} = P^T$) and a diagonal matrix $D$ such that

$$A = PDP^T = PDP^{-1} \tag{1}$$

Such a diagonalization requires $n$ linearly independent and orthonormal eigenvectors. When is this possible? If $A$ is orthogonally diagonalizable as in (1), then

$$A^T = (PDP^T)^T = P^{TT} D^T P^T = PDP^T = A$$

Thus $A$ is symmetric! Theorem 2 below shows that, conversely, every symmetric matrix is orthogonally diagonalizable. The proof is much harder and is omitted; the main idea for a proof will be given after Theorem 3.

**THEOREM 2**

An $n \times n$ matrix $A$ is orthogonally diagonalizable if and only if $A$ is a symmetric matrix.

This theorem is rather amazing, because the work in Chapter 5 would suggest that it is usually impossible to tell when a matrix is diagonalizable. But this is not the case for symmetric matrices.

The next example treats a matrix whose eigenvalues are not all distinct.

**EXAMPLE 3** Orthogonally diagonalize the matrix $A = \begin{bmatrix} 3 & -2 & 4 \\ -2 & 6 & 2 \\ 4 & 2 & 3 \end{bmatrix}$, whose characteristic equation is

$$0 = -\lambda^3 + 12\lambda^2 - 21\lambda - 98 = -(\lambda - 7)^2(\lambda + 2)$$

**SOLUTION** The usual calculations produce bases for the eigenspaces:

$$\lambda = 7: \mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} -1/2 \\ 1 \\ 0 \end{bmatrix}; \qquad \lambda = -2: \mathbf{v}_3 = \begin{bmatrix} -1 \\ -1/2 \\ 1 \end{bmatrix}$$

Although $\mathbf{v}_1$ and $\mathbf{v}_2$ are linearly independent, they are not orthogonal. Recall from Section 6.2 that the projection of $\mathbf{v}_2$ onto $\mathbf{v}_1$ is $\dfrac{\mathbf{v}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1$, and the component of $\mathbf{v}_2$ orthogonal to $\mathbf{v}_1$ is

$$\mathbf{z}_2 = \mathbf{v}_2 - \frac{\mathbf{v}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 = \begin{bmatrix} -1/2 \\ 1 \\ 0 \end{bmatrix} - \frac{-1/2}{2}\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -1/4 \\ 1 \\ 1/4 \end{bmatrix}$$

Then $\{\mathbf{v}_1, \mathbf{z}_2\}$ is an orthogonal set in the eigenspace for $\lambda = 7$. (Note that $\mathbf{z}_2$ is a linear combination of the eigenvectors $\mathbf{v}_1$ and $\mathbf{v}_2$, so $\mathbf{z}_2$ is in the eigenspace. This construction of $\mathbf{z}_2$ is just the Gram–Schmidt process of Section 6.4.) Since the eigenspace is two-dimensional (with basis $\mathbf{v}_1, \mathbf{v}_2$), the orthogonal set $\{\mathbf{v}_1, \mathbf{z}_2\}$ is an *orthogonal basis* for the eigenspace, by the Basis Theorem. (See Section 2.9 or 4.5.)

Normalize $\mathbf{v}_1$ and $\mathbf{z}_2$ to obtain the following orthonormal basis for the eigenspace for $\lambda = 7$:

$$\mathbf{u}_1 = \begin{bmatrix} 1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{bmatrix}, \quad \mathbf{u}_2 = \begin{bmatrix} -1/\sqrt{18} \\ 4/\sqrt{18} \\ 1/\sqrt{18} \end{bmatrix}$$

An orthonormal basis for the eigenspace for $\lambda = -2$ is

$$\mathbf{u}_3 = \frac{1}{\|2\mathbf{v}_3\|} 2\mathbf{v}_3 = \frac{1}{3}\begin{bmatrix} -2 \\ -1 \\ 2 \end{bmatrix} = \begin{bmatrix} -2/3 \\ -1/3 \\ 2/3 \end{bmatrix}$$

By Theorem 1, $\mathbf{u}_3$ is orthogonal to the other eigenvectors $\mathbf{u}_1$ and $\mathbf{u}_2$. Hence $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ is an orthonormal set. Let

$$P = [\,\mathbf{u}_1 \quad \mathbf{u}_2 \quad \mathbf{u}_3\,] = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{18} & -2/3 \\ 0 & 4/\sqrt{18} & -1/3 \\ 1/\sqrt{2} & 1/\sqrt{18} & 2/3 \end{bmatrix}, \quad D = \begin{bmatrix} 7 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & -2 \end{bmatrix}$$

Then $P$ orthogonally diagonalizes $A$, and $A = PDP^{-1}$. ■

In Example 3, the eigenvalue 7 has multiplicity two and the eigenspace is two-dimensional. This fact is not accidental, as the next theorem shows.

## The Spectral Theorem

The set of eigenvalues of a matrix $A$ is sometimes called the *spectrum* of $A$, and the following description of the eigenvalues is called a *spectral theorem.*

**THEOREM 3**

**The Spectral Theorem for Symmetric Matrices**

An $n \times n$ symmetric matrix $A$ has the following properties:

a. $A$ has $n$ real eigenvalues, counting multiplicities.

b. The dimension of the eigenspace for each eigenvalue $\lambda$ equals the multiplicity of $\lambda$ as a root of the characteristic equation.

c. The eigenspaces are mutually orthogonal, in the sense that eigenvectors corresponding to different eigenvalues are orthogonal.

d. $A$ is orthogonally diagonalizable.

Part (a) follows from Exercise 24 in Section 5.5. Part (b) follows easily from part (d). (See Exercise 31.) Part (c) is Theorem 1. Because of (a), a proof of (d) can be given using Exercise 32 and the Schur factorization discussed in Supplementary Exercise 16 in Chapter 6. The details are omitted.

## Spectral Decomposition

Suppose $A = PDP^{-1}$, where the columns of $P$ are orthonormal eigenvectors $\mathbf{u}_1, \dots, \mathbf{u}_n$ of $A$ and the corresponding eigenvalues $\lambda_1, \dots, \lambda_n$ are in the diagonal matrix $D$. Then, since $P^{-1} = P^T$,

$$A = PDP^T = \begin{bmatrix} \mathbf{u}_1 & \cdots & \mathbf{u}_n \end{bmatrix} \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix} \begin{bmatrix} \mathbf{u}_1^T \\ \vdots \\ \mathbf{u}_n^T \end{bmatrix}$$

$$= \begin{bmatrix} \lambda_1\mathbf{u}_1 & \cdots & \lambda_n\mathbf{u}_n \end{bmatrix} \begin{bmatrix} \mathbf{u}_1^T \\ \vdots \\ \mathbf{u}_n^T \end{bmatrix}$$

Using the column–row expansion of a product (Theorem 10 in Section 2.4), we can write

$$A = \lambda_1\mathbf{u}_1\mathbf{u}_1^T + \lambda_2\mathbf{u}_2\mathbf{u}_2^T + \cdots + \lambda_n\mathbf{u}_n\mathbf{u}_n^T \tag{2}$$

This representation of $A$ is called a **spectral decomposition** of $A$ because it breaks up $A$ into pieces determined by the spectrum (eigenvalues) of $A$. Each term in (2) is an $n \times n$ matrix of rank 1. For example, every column of $\lambda_1\mathbf{u}_1\mathbf{u}_1^T$ is a multiple of $\mathbf{u}_1$. Furthermore, each matrix $\mathbf{u}_j\mathbf{u}_j^T$ is a **projection matrix** in the sense that for each $\mathbf{x}$ in $\mathbb{R}^n$, the vector $(\mathbf{u}_j\mathbf{u}_j^T)\mathbf{x}$ is the orthogonal projection of $\mathbf{x}$ onto the subspace spanned by $\mathbf{u}_j$. (See Exercise 35.)

**EXAMPLE 4** Construct a spectral decomposition of the matrix $A$ that has the orthogonal diagonalization

$$A = \begin{bmatrix} 7 & 2 \\ 2 & 4 \end{bmatrix} = \begin{bmatrix} 2/\sqrt{5} & -1/\sqrt{5} \\ 1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix} \begin{bmatrix} 8 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 2/\sqrt{5} & 1/\sqrt{5} \\ -1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix}$$

**SOLUTION** Denote the columns of $P$ by $\mathbf{u}_1$ and $\mathbf{u}_2$. Then

$$A = 8\mathbf{u}_1\mathbf{u}_1^T + 3\mathbf{u}_2\mathbf{u}_2^T$$

To verify this decomposition of $A$, compute

$$\mathbf{u}_1\mathbf{u}_1^T = \begin{bmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix} \begin{bmatrix} 2/\sqrt{5} & 1/\sqrt{5} \end{bmatrix} = \begin{bmatrix} 4/5 & 2/5 \\ 2/5 & 1/5 \end{bmatrix}$$

$$\mathbf{u}_2\mathbf{u}_2^T = \begin{bmatrix} -1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix} \begin{bmatrix} -1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix} = \begin{bmatrix} 1/5 & -2/5 \\ -2/5 & 4/5 \end{bmatrix}$$

and

$$8\mathbf{u}_1\mathbf{u}_1^T + 3\mathbf{u}_2\mathbf{u}_2^T = \begin{bmatrix} 32/5 & 16/5 \\ 16/5 & 8/5 \end{bmatrix} + \begin{bmatrix} 3/5 & -6/5 \\ -6/5 & 12/5 \end{bmatrix} = \begin{bmatrix} 7 & 2 \\ 2 & 4 \end{bmatrix} = A \quad \blacksquare$$

### NUMERICAL NOTE

When $A$ is symmetric and not too large, modern high-performance computer algorithms calculate eigenvalues and eigenvectors with great precision. They apply a sequence of similarity transformations to $A$ involving orthogonal matrices. The diagonal entries of the transformed matrices converge rapidly to the eigenvalues of $A$. (See the Numerical Notes in Section 5.2.) Using orthogonal matrices generally prevents numerical errors from accumulating during the process. When $A$ is symmetric, the sequence of orthogonal matrices combines to form an orthogonal matrix whose columns are eigenvectors of $A$.

A nonsymmetric matrix cannot have a full set of orthogonal eigenvectors, but the algorithm still produces fairly accurate eigenvalues. After that, nonorthogonal techniques are needed to calculate eigenvectors.

### PRACTICE PROBLEMS

1. Show that if $A$ is a symmetric matrix, then $A^2$ is symmetric.
2. Show that if $A$ is orthogonally diagonalizable, then so is $A^2$.

## 7.1 EXERCISES

Determine which of the matrices in Exercises 1–6 are symmetric.

**1.** $\begin{bmatrix} 3 & 5 \\ 5 & -7 \end{bmatrix}$

**2.** $\begin{bmatrix} -3 & 5 \\ -5 & 3 \end{bmatrix}$

**3.** $\begin{bmatrix} 2 & 2 \\ 4 & 4 \end{bmatrix}$

**4.** $\begin{bmatrix} 0 & 8 & 3 \\ 8 & 0 & -2 \\ 3 & -2 & 0 \end{bmatrix}$

**5.** $\begin{bmatrix} -6 & 2 & 0 \\ 0 & -6 & 2 \\ 0 & 0 & -6 \end{bmatrix}$

**6.** $\begin{bmatrix} 1 & 2 & 1 & 2 \\ 2 & 1 & 2 & 1 \\ 1 & 2 & 1 & 2 \end{bmatrix}$

Determine which of the matrices in Exercises 7–12 are orthogonal. If orthogonal, find the inverse.

**7.** $\begin{bmatrix} .6 & .8 \\ .8 & -.6 \end{bmatrix}$

**8.** $\begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$

**9.** $\begin{bmatrix} -5 & 2 \\ 2 & 5 \end{bmatrix}$

**10.** $\begin{bmatrix} -1 & 2 & 2 \\ 2 & -1 & 2 \\ 2 & 2 & -1 \end{bmatrix}$

**11.** $\begin{bmatrix} 2/3 & 2/3 & 1/3 \\ 0 & 1/\sqrt{5} & -2/\sqrt{5} \\ \sqrt{5}/3 & -4/\sqrt{45} & -2/\sqrt{45} \end{bmatrix}$

**12.** $\begin{bmatrix} .5 & .5 & -.5 & -.5 \\ -.5 & .5 & -.5 & .5 \\ .5 & .5 & .5 & .5 \\ -.5 & .5 & .5 & -.5 \end{bmatrix}$

Orthogonally diagonalize the matrices in Exercises 13–22, giving an orthogonal matrix $P$ and a diagonal matrix $D$. To save you time, the eigenvalues in Exercises 17–22 are: (17) 5, 2, $-2$; (18) 25, 3, $-50$; (19) 7, $-2$; (20) 13, 7, 1; (21) 9, 5, 1; (22) 2, 0.

**13.** $\begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$

**14.** $\begin{bmatrix} 1 & 5 \\ 5 & 1 \end{bmatrix}$

**15.** $\begin{bmatrix} 16 & -4 \\ -4 & 1 \end{bmatrix}$

**16.** $\begin{bmatrix} -7 & 24 \\ 24 & 7 \end{bmatrix}$

**17.** $\begin{bmatrix} 1 & 1 & 3 \\ 1 & 3 & 1 \\ 3 & 1 & 1 \end{bmatrix}$

**18.** $\begin{bmatrix} -2 & -36 & 0 \\ -36 & -23 & 0 \\ 0 & 0 & 3 \end{bmatrix}$

**19.** $\begin{bmatrix} 3 & -2 & 4 \\ -2 & 6 & 2 \\ 4 & 2 & 3 \end{bmatrix}$

**20.** $\begin{bmatrix} 7 & -4 & 4 \\ -4 & 5 & 0 \\ 4 & 0 & 9 \end{bmatrix}$

**21.** $\begin{bmatrix} 4 & 1 & 3 & 1 \\ 1 & 4 & 1 & 3 \\ 3 & 1 & 4 & 1 \\ 1 & 3 & 1 & 4 \end{bmatrix}$

**22.** $\begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 2 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}$

**23.** Let $A = \begin{bmatrix} 3 & 1 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 3 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$. Verify that 2 is an eigenvalue of $A$ and $\mathbf{v}$ is an eigenvector. Then orthogonally diagonalize $A$.

**24.** Let $A = \begin{bmatrix} 5 & -4 & -2 \\ -4 & 5 & 2 \\ -2 & 2 & 2 \end{bmatrix}$, $\mathbf{v}_1 = \begin{bmatrix} -2 \\ 2 \\ 1 \end{bmatrix}$, and $\mathbf{v}_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$. Verify that $\mathbf{v}_1$ and $\mathbf{v}_2$ are eigenvectors of $A$. Then orthogonally diagonalize $A$.

In Exercises 25 and 26, mark each statement True or False. Justify each answer.

**25.** a. An $n \times n$ matrix that is orthogonally diagonalizable must be symmetric.

b. If $A^T = A$ and if vectors $\mathbf{u}$ and $\mathbf{v}$ satisfy $A\mathbf{u} = 3\mathbf{u}$ and $A\mathbf{v} = 4\mathbf{v}$, then $\mathbf{u}\cdot\mathbf{v} = 0$.

c. An $n \times n$ symmetric matrix has $n$ distinct real eigenvalues.

d. For a nonzero $\mathbf{v}$ in $\mathbb{R}^n$, the matrix $\mathbf{v}\mathbf{v}^T$ is called a projection matrix.

**26.** a. Every symmetric matrix is orthogonally diagonalizable.

b. If $B = PDP^T$, where $P^T = P^{-1}$ and $D$ is a diagonal matrix, then $B$ is a symmetric matrix.

c. An orthogonal matrix is orthogonally diagonalizable.

d. The dimension of an eigenspace of a symmetric matrix equals the multiplicity of the corresponding eigenvalue.

**27.** Suppose $A$ is a symmetric $n \times n$ matrix and $B$ is any $n \times m$ matrix. Show that $B^TAB$, $B^TB$, and $BB^T$ are symmetric matrices.

**28.** Show that if $A$ is an $n \times n$ symmetric matrix, then $(A\mathbf{x})\cdot\mathbf{y} = \mathbf{x}\cdot(A\mathbf{y})$ for all $\mathbf{x}, \mathbf{y}$ in $\mathbb{R}^n$.

**29.** Suppose $A$ is invertible and orthogonally diagonalizable. Explain why $A^{-1}$ is also orthogonally diagonalizable.

**30.** Suppose $A$ and $B$ are both orthogonally diagonalizable and $AB = BA$. Explain why $AB$ is also orthogonally diagonalizable.

**31.** Let $A = PDP^{-1}$, where $P$ is orthogonal and $D$ is diagonal, and let $\lambda$ be an eigenvalue of $A$ of multiplicity $k$. Then $\lambda$ appears $k$ times on the diagonal of $D$. Explain why the dimension of the eigenspace for $\lambda$ is $k$.

**32.** Suppose $A = PRP^{-1}$, where $P$ is orthogonal and $R$ is upper triangular. Show that if $A$ is symmetric, then $R$ is symmetric and hence is actually a diagonal matrix.

**33.** Construct a spectral decomposition of $A$ from Example 2.

**34.** Construct a spectral decomposition of $A$ from Example 3.

**35.** Let $\mathbf{u}$ be a unit vector in $\mathbb{R}^n$, and let $B = \mathbf{u}\mathbf{u}^T$.

a. Given any $\mathbf{x}$ in $\mathbb{R}^n$, compute $B\mathbf{x}$ and show that $B\mathbf{x}$ is the orthogonal projection of $\mathbf{x}$ onto $\mathbf{u}$, as described in Section 6.2.

b. Show that $B$ is a symmetric matrix and $B^2 = B$.

c. Show that $\mathbf{u}$ is an eigenvector of $B$. What is the corresponding eigenvalue?

**36.** Let $B$ be an $n \times n$ symmetric matrix such that $B^2 = B$. Any such matrix is called a **projection matrix** (or an **orthogonal projection matrix**). Given any $\mathbf{y}$ in $\mathbb{R}^n$, let $\hat{\mathbf{y}} = B\mathbf{y}$ and $\mathbf{z} = \mathbf{y} - \hat{\mathbf{y}}$.

a. Show that $\mathbf{z}$ is orthogonal to $\hat{\mathbf{y}}$.

b. Let $W$ be the column space of $B$. Show that $\mathbf{y}$ is the sum of a vector in $W$ and a vector in $W^\perp$. Why does this prove that $B\mathbf{y}$ is the orthogonal projection of $\mathbf{y}$ onto the column space of $B$?

[**M**] Orthogonally diagonalize the matrices in Exercises 37–40. To practice the methods of this section, do not use an eigenvector routine from your matrix program. Instead, use the program to find the eigenvalues, and, for each eigenvalue $\lambda$, find an orthonormal basis for $\text{Nul}(A - \lambda I)$, as in Examples 2 and 3.

**37.** $\begin{bmatrix} 5 & 2 & 9 & -6 \\ 2 & 5 & -6 & 9 \\ 9 & -6 & 5 & 2 \\ -6 & 9 & 2 & 5 \end{bmatrix}$

**38.** $\begin{bmatrix} .38 & -.18 & -.06 & -.04 \\ -.18 & .59 & -.04 & .12 \\ -.06 & -.04 & .47 & -.12 \\ -.04 & .12 & -.12 & .41 \end{bmatrix}$

**39.** $\begin{bmatrix} .31 & .58 & .08 & .44 \\ .58 & -.56 & .44 & -.58 \\ .08 & .44 & .19 & -.08 \\ .44 & -.58 & -.08 & .31 \end{bmatrix}$

**40.** $\begin{bmatrix} 10 & 2 & 2 & -6 & 9 \\ 2 & 10 & 2 & -6 & 9 \\ 2 & 2 & 10 & -6 & 9 \\ -6 & -6 & -6 & 26 & 9 \\ 9 & 9 & 9 & 9 & -19 \end{bmatrix}$

## SOLUTIONS TO PRACTICE PROBLEMS

**1.** $(A^2)^T = (AA)^T = A^TA^T$, by a property of transposes. By hypothesis, $A^T = A$. So $(A^2)^T = AA = A^2$, which shows that $A^2$ is symmetric.

**2.** If $A$ is orthogonally diagonalizable, then $A$ is symmetric, by Theorem 2. By Practice Problem 1, $A^2$ is symmetric and hence is orthogonally diagonalizable (Theorem 2).

## 7.2 QUADRATIC FORMS

Until now, our attention in this text has focused on linear equations, except for the sums of squares encountered in Chapter 6 when computing $\mathbf{x}^T\mathbf{x}$. Such sums and more general expressions, called *quadratic forms*, occur frequently in applications of linear algebra to engineering (in design criteria and optimization) and signal processing (as output noise power). They also arise, for example, in physics (as potential and kinetic energy), differential geometry (as normal curvature of surfaces), economics (as utility functions), and statistics (in confidence ellipsoids). Some of the mathematical background for such applications flows easily from our work on symmetric matrices.

A **quadratic form** on $\mathbb{R}^n$ is a function $Q$ defined on $\mathbb{R}^n$ whose value at a vector $\mathbf{x}$ in $\mathbb{R}^n$ can be computed by an expression of the form $Q(\mathbf{x}) = \mathbf{x}^TA\mathbf{x}$, where $A$ is an $n \times n$ symmetric matrix. The matrix $A$ is called the **matrix of the quadratic form**.

The simplest example of a nonzero quadratic form is $Q(\mathbf{x}) = \mathbf{x}^TI\mathbf{x} = \|\mathbf{x}\|^2$. Examples 1 and 2 show the connection between any symmetric matrix $A$ and the quadratic form $\mathbf{x}^TA\mathbf{x}$.

**EXAMPLE 1** Let $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$. Compute $\mathbf{x}^TA\mathbf{x}$ for the following matrices:

a. $A = \begin{bmatrix} 4 & 0 \\ 0 & 3 \end{bmatrix}$ b. $A = \begin{bmatrix} 3 & -2 \\ -2 & 7 \end{bmatrix}$

**SOLUTION**

a. $\mathbf{x}^TA\mathbf{x} = [\,x_1 \quad x_2\,]\begin{bmatrix} 4 & 0 \\ 0 & 3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = [\,x_1 \quad x_2\,]\begin{bmatrix} 4x_1 \\ 3x_2 \end{bmatrix} = 4x_1^2 + 3x_2^2.$

b. There are two $-2$ entries in $A$. Watch how they enter the calculations. The $(1, 2)$-entry in $A$ is in boldface type.

$$\begin{aligned}
\mathbf{x}^TA\mathbf{x} &= [\,x_1 \quad x_2\,]\begin{bmatrix} 3 & \mathbf{-2} \\ -2 & 7 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = [\,x_1 \quad x_2\,]\begin{bmatrix} 3x_1 - \mathbf{2}x_2 \\ -2x_1 + 7x_2 \end{bmatrix} \\
&= x_1(3x_1 - \mathbf{2}x_2) + x_2(-2x_1 + 7x_2) \\
&= 3x_1^2 - \mathbf{2}x_1x_2 - 2x_2x_1 + 7x_2^2 \\
&= 3x_1^2 - 4x_1x_2 + 7x_2^2
\end{aligned}$$

∎

The presence of $-4x_1x_2$ in the quadratic form in Example 1(b) is due to the $-2$ entries off the diagonal in the matrix $A$. In contrast, the quadratic form associated with the diagonal matrix $A$ in Example 1(a) has no $x_1x_2$ *cross-product* term.

**EXAMPLE 2** For $\mathbf{x}$ in $\mathbb{R}^3$, let $Q(\mathbf{x}) = 5x_1^2 + 3x_2^2 + 2x_3^2 - x_1x_2 + 8x_2x_3$. Write this quadratic form as $\mathbf{x}^TA\mathbf{x}$.

**SOLUTION** The coefficients of $x_1^2$, $x_2^2$, $x_3^2$ go on the diagonal of $A$. To make $A$ symmetric, the coefficient of $x_ix_j$ for $i \neq j$ must be split evenly between the $(i, j)$- and $(j, i)$-entries in $A$. The coefficient of $x_1x_3$ is 0. It is readily checked that

$$Q(\mathbf{x}) = \mathbf{x}^TA\mathbf{x} = [\,x_1 \quad x_2 \quad x_3\,]\begin{bmatrix} 5 & -1/2 & 0 \\ -1/2 & 3 & 4 \\ 0 & 4 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

∎

**EXAMPLE 3** Let $Q(\mathbf{x}) = x_1^2 - 8x_1x_2 - 5x_2^2$. Compute the value of $Q(\mathbf{x})$ for $\mathbf{x} = \begin{bmatrix} -3 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ -2 \end{bmatrix}$, and $\begin{bmatrix} 1 \\ -3 \end{bmatrix}$.

**SOLUTION**

$$Q(-3, 1) = (-3)^2 - 8(-3)(1) - 5(1)^2 = 28$$
$$Q(2, -2) = (2)^2 - 8(2)(-2) - 5(-2)^2 = 16$$
$$Q(1, -3) = (1)^2 - 8(1)(-3) - 5(-3)^2 = -20$$

In some cases, quadratic forms are easier to use when they have no cross-product terms—that is, when the matrix of the quadratic form is a diagonal matrix. Fortunately, the cross-product term can be eliminated by making a suitable change of variable.

## Change of Variable in a Quadratic Form

If $\mathbf{x}$ represents a variable vector in $\mathbb{R}^n$, then a **change of variable** is an equation of the form

$$\mathbf{x} = P\mathbf{y}, \qquad \text{or equivalently,} \qquad \mathbf{y} = P^{-1}\mathbf{x} \tag{1}$$

where $P$ is an invertible matrix and $\mathbf{y}$ is a new variable vector in $\mathbb{R}^n$. Here $\mathbf{y}$ is the coordinate vector of $\mathbf{x}$ relative to the basis of $\mathbb{R}^n$ determined by the columns of $P$. (See Section 4.4.)

If the change of variable (1) is made in a quadratic form $\mathbf{x}^TA\mathbf{x}$, then

$$\mathbf{x}^TA\mathbf{x} = (P\mathbf{y})^TA(P\mathbf{y}) = \mathbf{y}^TP^TAP\mathbf{y} = \mathbf{y}^T(P^TAP)\mathbf{y} \tag{2}$$

and the new matrix of the quadratic form is $P^TAP$. Since $A$ is symmetric, Theorem 2 guarantees that there is an *orthogonal* matrix $P$ such that $P^TAP$ is a diagonal matrix $D$, and the quadratic form in (2) becomes $\mathbf{y}^TD\mathbf{y}$. This is the strategy of the next example.

**EXAMPLE 4** Make a change of variable that transforms the quadratic form in Example 3 into a quadratic form with no cross-product term.

**SOLUTION** The matrix of the quadratic form in Example 3 is

$$A = \begin{bmatrix} 1 & -4 \\ -4 & -5 \end{bmatrix}$$

The first step is to orthogonally diagonalize $A$. Its eigenvalues turn out to be $\lambda = 3$ and $\lambda = -7$. Associated unit eigenvectors are

$$\lambda = 3: \begin{bmatrix} 2/\sqrt{5} \\ -1/\sqrt{5} \end{bmatrix}; \qquad \lambda = -7: \begin{bmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix}$$

These vectors are automatically orthogonal (because they correspond to distinct eigenvalues) and so provide an orthonormal basis for $\mathbb{R}^2$. Let

$$P = \begin{bmatrix} 2/\sqrt{5} & 1/\sqrt{5} \\ -1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix}, \qquad D = \begin{bmatrix} 3 & 0 \\ 0 & -7 \end{bmatrix}$$

Then $A = PDP^{-1}$ and $D = P^{-1}AP = P^TAP$, as pointed out earlier. A suitable change of variable is

$$\mathbf{x} = P\mathbf{y}, \qquad \text{where } \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad \text{and} \quad \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

Then

$$\begin{aligned} x_1^2 - 8x_1x_2 - 5x_2^2 &= \mathbf{x}^T A\mathbf{x} = (P\mathbf{y})^T A(P\mathbf{y}) \\ &= \mathbf{y}^T P^T A P\mathbf{y} = \mathbf{y}^T D\mathbf{y} \\ &= 3y_1^2 - 7y_2^2 \end{aligned}$$

To illustrate the meaning of the equality of quadratic forms in Example 4, we can compute $Q(\mathbf{x})$ for $\mathbf{x} = (2, -2)$ using the new quadratic form. First, since $\mathbf{x} = P\mathbf{y}$,

$$\mathbf{y} = P^{-1}\mathbf{x} = P^T\mathbf{x}$$

so

$$\mathbf{y} = \begin{bmatrix} 2/\sqrt{5} & -1/\sqrt{5} \\ 1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix} \begin{bmatrix} 2 \\ -2 \end{bmatrix} = \begin{bmatrix} 6/\sqrt{5} \\ -2/\sqrt{5} \end{bmatrix}$$

Hence

$$\begin{aligned} 3y_1^2 - 7y_2^2 &= 3(6/\sqrt{5})^2 - 7(-2/\sqrt{5})^2 = 3(36/5) - 7(4/5) \\ &= 80/5 = 16 \end{aligned}$$

This is the value of $Q(\mathbf{x})$ in Example 3 when $\mathbf{x} = (2, -2)$. See Fig. 1.

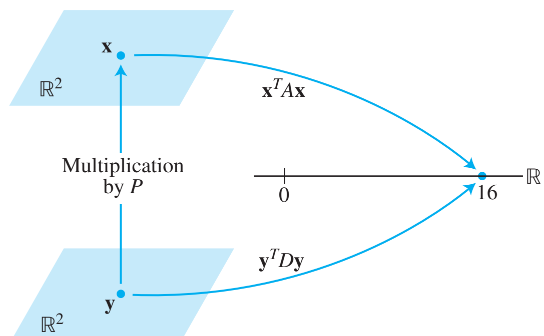



**FIGURE 1** Change of variable in $\mathbf{x}^T A\mathbf{x}$.

Example 4 illustrates the following theorem. The proof of the theorem was essentially given before Example 4.

**THEOREM 4**

**The Principal Axes Theorem**

Let $A$ be an $n \times n$ symmetric matrix. Then there is an orthogonal change of variable, $\mathbf{x} = P\mathbf{y}$, that transforms the quadratic form $\mathbf{x}^T A\mathbf{x}$ into a quadratic form $\mathbf{y}^T D\mathbf{y}$ with no cross-product term.

The columns of $P$ in the theorem are called the **principal axes** of the quadratic form $\mathbf{x}^T A\mathbf{x}$. The vector $\mathbf{y}$ is the coordinate vector of $\mathbf{x}$ relative to the orthonormal basis of $\mathbb{R}^n$ given by these principal axes.

## A Geometric View of Principal Axes

Suppose $Q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$, where $A$ is an invertible $2 \times 2$ symmetric matrix, and let $c$ be a constant. It can be shown that the set of all $\mathbf{x}$ in $\mathbb{R}^2$ that satisfy

$$\mathbf{x}^T A\mathbf{x} = c \tag{3}$$

either corresponds to an ellipse (or circle), a hyperbola, two intersecting lines, or a single point, or contains no points at all. If $A$ is a diagonal matrix, the graph is in *standard position*, such as in Fig. 2. If $A$ is not a diagonal matrix, the graph of equation (3) is

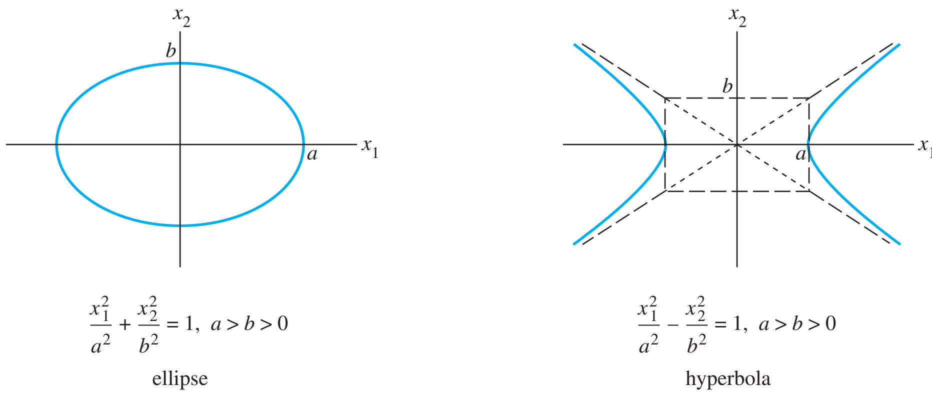



**FIGURE 2** An ellipse and a hyperbola in standard position.

rotated out of standard position, as in Fig. 3. Finding the *principal axes* (determined by the eigenvectors of $A$) amounts to finding a new coordinate system with respect to which the graph is in standard position.

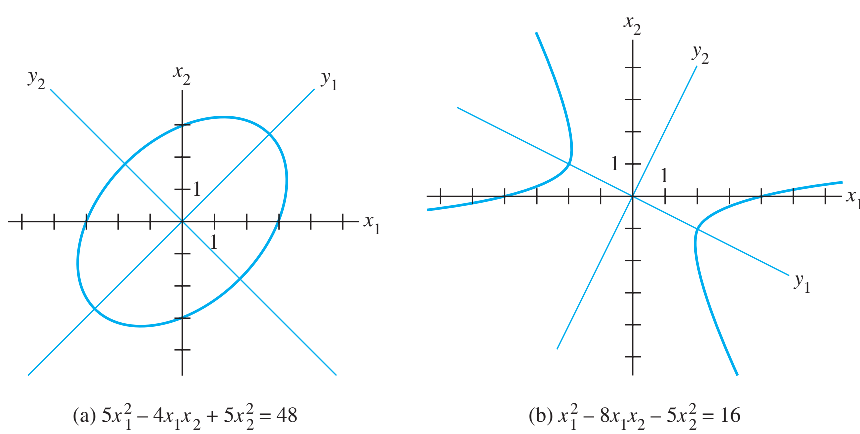



**FIGURE 3** An ellipse and a hyperbola *not* in standard position.

The hyperbola in Fig. 3(b) is the graph of the equation $\mathbf{x}^TA\mathbf{x} = 16$, where $A$ is the matrix in Example 4. The positive $y_1$-axis in Fig. 3(b) is in the direction of the first column of the matrix $P$ in Example 4, and the positive $y_2$-axis is in the direction of the second column of $P$.

**EXAMPLE 5** The ellipse in Fig. 3(a) is the graph of the equation $5x_1^2 - 4x_1x_2 + 5x_2^2 = 48$. Find a change of variable that removes the cross-product term from the equation.

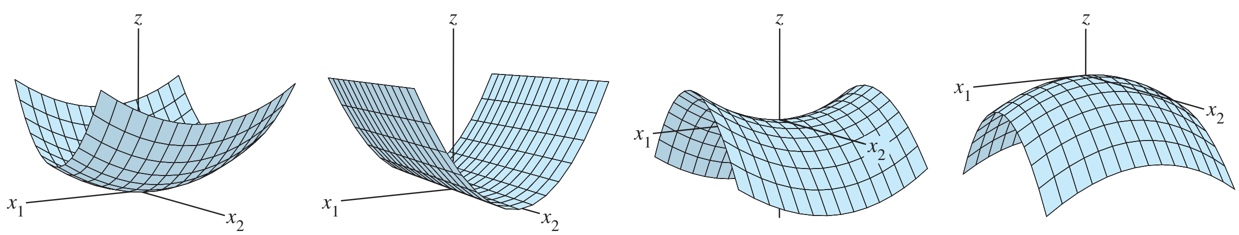
**SOLUTION** The matrix of the quadratic form is $A = \begin{bmatrix} 5 & -2 \\ -2 & 5 \end{bmatrix}$. The eigenvalues of $A$ turn out to be 3 and 7, with corresponding unit eigenvectors

$$\mathbf{u}_1 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}, \quad \mathbf{u}_2 = \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$$

Let $P = [\,\mathbf{u}_1 \quad \mathbf{u}_2\,] = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$. Then $P$ orthogonally diagonalizes $A$, so the change of variable $\mathbf{x} = P\mathbf{y}$ produces the quadratic form $\mathbf{y}^T D\mathbf{y} = 3y_1^2 + 7y_2^2$. The new axes for this change of variable are shown in Fig. 3(a). ■

## Classifying Quadratic Forms

When $A$ is an $n \times n$ matrix, the quadratic form $Q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ is a real-valued function with domain $\mathbb{R}^n$. Figure 4 displays the graphs of four quadratic forms with domain $\mathbb{R}^2$. For each point $\mathbf{x} = (x_1, x_2)$ in the domain of a quadratic form $Q$, the graph displays the point $(x_1, x_2, z)$ where $z = Q(\mathbf{x})$. Notice that except at $\mathbf{x} = \mathbf{0}$, the values of $Q(\mathbf{x})$ are all positive in Fig. 4(a) and all negative in Fig. 4(d). The horizontal cross-sections of the graphs are ellipses in Figs. 4(a) and 4(d) and hyperbolas in Fig. 4(c).




(a) $z = 3x_1^2 + 7x_2^2$ (b) $z = 3x_1^2$ (c) $z = 3x_1^2 - 7x_2^2$ (d) $z = -3x_1^2 - 7x_2^2$

**FIGURE 4** Graphs of quadratic forms.

The simple $2 \times 2$ examples in Fig. 4 illustrate the following definitions.

**DEFINITION**

A quadratic form $Q$ is:

a. **positive definite** if $Q(\mathbf{x}) > 0$ for all $\mathbf{x} \neq \mathbf{0}$,

b. **negative definite** if $Q(\mathbf{x}) < 0$ for all $\mathbf{x} \neq \mathbf{0}$,

c. **indefinite** if $Q(\mathbf{x})$ assumes both positive and negative values.

Also, $Q$ is said to be **positive semidefinite** if $Q(\mathbf{x}) \geq 0$ for all $\mathbf{x}$, and to be **negative semidefinite** if $Q(\mathbf{x}) \leq 0$ for all $\mathbf{x}$. The quadratic forms in parts (a) and (b) of Fig. 4 are both positive semidefinite, but the form in (a) is better described as positive definite.

Theorem 5 characterizes some quadratic forms in terms of eigenvalues.

**THEOREM 5**

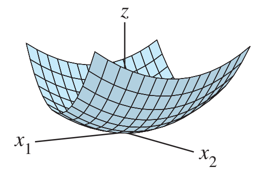
**Quadratic Forms and Eigenvalues**

Let $A$ be an $n \times n$ symmetric matrix. Then a quadratic form $\mathbf{x}^T A\mathbf{x}$ is:

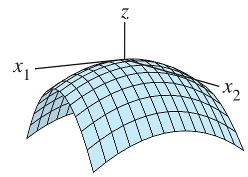
a. positive definite if and only if the eigenvalues of $A$ are all positive,

b. negative definite if and only if the eigenvalues of $A$ are all negative, or

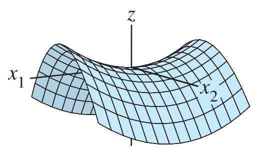
c. indefinite if and only if $A$ has both positive and negative eigenvalues.



Positive definite



Negative definite



Indefinite

**PROOF** By the Principal Axes Theorem, there exists an orthogonal change of variable $\mathbf{x} = P\mathbf{y}$ such that

$$Q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x} = \mathbf{y}^T D\mathbf{y} = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2 \tag{4}$$

where $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of $A$. Since $P$ is invertible, there is a one-to-one correspondence between all nonzero $\mathbf{x}$ and all nonzero $\mathbf{y}$. Thus the values of $Q(\mathbf{x})$ for $\mathbf{x} \neq \mathbf{0}$ coincide with the values of the expression on the right side of (4), which is obviously controlled by the signs of the eigenvalues $\lambda_1, \ldots, \lambda_n$, in the three ways described in the theorem. ■

**EXAMPLE 6** Is $Q(\mathbf{x}) = 3x_1^2 + 2x_2^2 + x_3^2 + 4x_1x_2 + 4x_2x_3$ positive definite?

**SOLUTION** Because of all the plus signs, this form "looks" positive definite. But the matrix of the form is

$$A = \begin{bmatrix} 3 & 2 & 0 \\ 2 & 2 & 2 \\ 0 & 2 & 1 \end{bmatrix}$$

and the eigenvalues of $A$ turn out to be 5, 2, and $-1$. So $Q$ is an indefinite quadratic form, not positive definite. ■

The classification of a quadratic form is often carried over to the matrix of the form. Thus a **positive definite matrix** $A$ is a *symmetric* matrix for which the quadratic form $\mathbf{x}^T A\mathbf{x}$ is positive definite. Other terms, such as **positive semidefinite matrix**, are defined analogously.

WEB

**NUMERICAL NOTE**

A fast way to determine whether a symmetric matrix $A$ is positive definite is to attempt to factor $A$ in the form $A = R^T R$, where $R$ is upper triangular with positive diagonal entries. (A slightly modified algorithm for an LU factorization is one approach.) Such a *Cholesky factorization* is possible if and only if $A$ is positive definite. See Supplementary Exercise 7 at the end of Chapter 7.

## PRACTICE PROBLEM

Describe a positive semidefinite matrix $A$ in terms of its eigenvalues.

WEB

## 7.2 EXERCISES

1. Compute the quadratic form $\mathbf{x}^T A\mathbf{x}$, when $A = \begin{bmatrix} 5 & 1/3 \\ 1/3 & 1 \end{bmatrix}$ and

   a. $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ b. $\mathbf{x} = \begin{bmatrix} 6 \\ 1 \end{bmatrix}$ c. $\mathbf{x} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$

2. Compute the quadratic form $\mathbf{x}^T A\mathbf{x}$, for $A = \begin{bmatrix} 4 & 3 & 0 \\ 3 & 2 & 1 \\ 0 & 1 & 1 \end{bmatrix}$ and

   a. $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ b. $\mathbf{x} = \begin{bmatrix} 2 \\ -1 \\ 5 \end{bmatrix}$ c. $\mathbf{x} = \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}$

3. Find the matrix of the quadratic form. Assume $\mathbf{x}$ is in $\mathbb{R}^2$.

   a. $10x_1^2 - 6x_1x_2 - 3x_2^2$ b. $5x_1^2 + 3x_1x_2$

4. Find the matrix of the quadratic form. Assume $\mathbf{x}$ is in $\mathbb{R}^2$.

   a. $20x_1^2 + 15x_1x_2 - 10x_2^2$ b. $x_1x_2$

**5.** Find the matrix of the quadratic form. Assume $\mathbf{x}$ is in $\mathbb{R}^3$.

a. $8x_1^2 + 7x_2^2 - 3x_3^2 - 6x_1x_2 + 4x_1x_3 - 2x_2x_3$

b. $4x_1x_2 + 6x_1x_3 - 8x_2x_3$

**6.** Find the matrix of the quadratic form. Assume $\mathbf{x}$ is in $\mathbb{R}^3$.

a. $5x_1^2 - x_2^2 + 7x_3^2 + 5x_1x_2 - 3x_1x_3$

b. $x_3^2 - 4x_1x_2 + 4x_2x_3$

**7.** Make a change of variable, $\mathbf{x} = P\mathbf{y}$, that transforms the quadratic form $x_1^2 + 10x_1x_2 + x_2^2$ into a quadratic form with no cross-product term. Give $P$ and the new quadratic form.

**8.** Let $A$ be the matrix of the quadratic form

$$9x_1^2 + 7x_2^2 + 11x_3^2 - 8x_1x_2 + 8x_1x_3$$

It can be shown that the eigenvalues of $A$ are 3, 9, and 15. Find an orthogonal matrix $P$ such that the change of variable $\mathbf{x} = P\mathbf{y}$ transforms $\mathbf{x}^T A\mathbf{x}$ into a quadratic form with no cross-product term. Give $P$ and the new quadratic form.

Classify the quadratic forms in Exercises 9–18. Then make a change of variable, $\mathbf{x} = P\mathbf{y}$, that transforms the quadratic form into one with no cross-product term. Write the new quadratic form. Construct $P$ using the methods of Section 7.1.

**9.** $3x_1^2 - 4x_1x_2 + 6x_2^2$

**10.** $9x_1^2 - 8x_1x_2 + 3x_2^2$

**11.** $2x_1^2 + 10x_1x_2 + 2x_2^2$

**12.** $-5x_1^2 + 4x_1x_2 - 2x_2^2$

**13.** $x_1^2 - 6x_1x_2 + 9x_2^2$

**14.** $8x_1^2 + 6x_1x_2$

**15.** **[M]** $-2x_1^2 - 6x_2^2 - 9x_3^2 - 9x_4^2 + 4x_1x_2 + 4x_1x_3 + 4x_1x_4 + 6x_3x_4$

**16.** **[M]** $4x_1^2 + 4x_2^2 + 4x_3^2 + 4x_4^2 + 3x_1x_2 + 3x_3x_4 - 4x_1x_4 + 4x_2x_3$

**17.** **[M]** $x_1^2 + x_2^2 + x_3^2 + x_4^2 + 9x_1x_2 - 12x_1x_4 + 12x_2x_3 + 9x_3x_4$

**18.** **[M]** $11x_1^2 - x_2^2 - 12x_1x_2 - 12x_1x_3 - 12x_1x_4 - 2x_3x_4$

**19.** What is the largest possible value of the quadratic form $5x_1^2 + 8x_2^2$ if $\mathbf{x} = (x_1, x_2)$ and $\mathbf{x}^T\mathbf{x} = 1$, that is, if $x_1^2 + x_2^2 = 1$? (Try some examples of $\mathbf{x}$.)

**20.** What is the largest value of the quadratic form $5x_1^2 - 3x_2^2$ if $\mathbf{x}^T\mathbf{x} = 1$?

In Exercises 21 and 22, matrices are $n \times n$ and vectors are in $\mathbb{R}^n$. Mark each statement True or False. Justify each answer.

**21.** a. The matrix of a quadratic form is a symmetric matrix.

b. A quadratic form has no cross-product terms if and only if the matrix of the quadratic form is a diagonal matrix.

c. The principal axes of a quadratic form $\mathbf{x}^T A\mathbf{x}$ are eigenvectors of $A$.

d. A positive definite quadratic form $Q$ satisfies $Q(\mathbf{x}) > 0$ for all $\mathbf{x}$ in $\mathbb{R}^n$.

e. If the eigenvalues of a symmetric matrix $A$ are all positive, then the quadratic form $\mathbf{x}^T A\mathbf{x}$ is positive definite.

f. A Cholesky factorization of a symmetric matrix $A$ has the form $A = R^T R$, for an upper triangular matrix $R$ with positive diagonal entries.

**22.** a. The expression $\|\mathbf{x}\|^2$ is a quadratic form.

b. If $A$ is symmetric and $P$ is an orthogonal matrix, then the change of variable $\mathbf{x} = P\mathbf{y}$ transforms $\mathbf{x}^T A\mathbf{x}$ into a quadratic form with no cross-product term.

c. If $A$ is a $2 \times 2$ symmetric matrix, then the set of $\mathbf{x}$ such that $\mathbf{x}^T A\mathbf{x} = c$ (for a constant $c$) corresponds to either a circle, an ellipse, or a hyperbola.

d. An indefinite quadratic form is either positive semidefinite or negative semidefinite.

e. If $A$ is symmetric and the quadratic form $\mathbf{x}^T A\mathbf{x}$ has only negative values for $\mathbf{x} \neq \mathbf{0}$, then the eigenvalues of $A$ are all negative.

Exercises 23 and 24 show how to classify a quadratic form $Q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$, when $A = \begin{bmatrix} a & b \\ b & d \end{bmatrix}$ and $\det A \neq 0$, without finding the eigenvalues of $A$.

**23.** If $\lambda_1$ and $\lambda_2$ are the eigenvalues of $A$, then the characteristic polynomial of $A$ can be written in two ways: $\det(A - \lambda I)$ and $(\lambda - \lambda_1)(\lambda - \lambda_2)$. Use this fact to show that $\lambda_1 + \lambda_2 = a + d$ (the diagonal entries of $A$) and $\lambda_1\lambda_2 = \det A$.

**24.** Verify the following statements.

a. $Q$ is positive definite if $\det A > 0$ and $a > 0$.

b. $Q$ is negative definite if $\det A > 0$ and $a < 0$.

c. $Q$ is indefinite if $\det A < 0$.

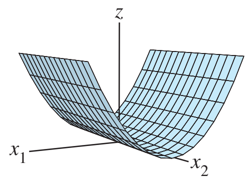
**25.** Show that if $B$ is $m \times n$, then $B^T B$ is positive semidefinite; and if $B$ is $n \times n$ and invertible, then $B^T B$ is positive definite.

**26.** Show that if an $n \times n$ matrix $A$ is positive definite, then there exists a positive definite matrix $B$ such that $A = B^T B$. [*Hint:* Write $A = PDP^T$, with $P^T = P^{-1}$. Produce a diagonal matrix $C$ such that $D = C^T C$, and let $B = PCP^T$. Show that $B$ works.]

**27.** Let $A$ and $B$ be symmetric $n \times n$ matrices whose eigenvalues are all positive. Show that the eigenvalues of $A + B$ are all positive. [*Hint:* Consider quadratic forms.]

**28.** Let $A$ be an $n \times n$ invertible symmetric matrix. Show that if the quadratic form $\mathbf{x}^T A\mathbf{x}$ is positive definite, then so is the quadratic form $\mathbf{x}^T A^{-1}\mathbf{x}$. [*Hint:* Consider eigenvalues.]

SG Mastering: Diagonalization and Quadratic Forms 7–7



Positive semidefinite

### SOLUTION TO PRACTICE PROBLEM

Make an orthogonal change of variable $\mathbf{x} = P\mathbf{y}$, and write

$$\mathbf{x}^T A\mathbf{x} = \mathbf{y}^T D\mathbf{y} = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2$$

as in equation (4). If an eigenvalue—say, $\lambda_i$—were negative, then $\mathbf{x}^T A\mathbf{x}$ would be negative for the $\mathbf{x}$ corresponding to $\mathbf{y} = \mathbf{e}_i$ (the $i$th column of $I_n$). So the eigenvalues of a positive semidefinite quadratic form must all be nonnegative. Conversely, if the eigenvalues are nonnegative, the expansion above shows that $\mathbf{x}^T A\mathbf{x}$ must be positive semidefinite.

## 7.3 CONSTRAINED OPTIMIZATION

Engineers, economists, scientists, and mathematicians often need to find the maximum or minimum value of a quadratic form $Q(\mathbf{x})$ for $\mathbf{x}$ in some specified set. Typically, the problem can be arranged so that $\mathbf{x}$ varies over the set of unit vectors. This *constrained optimization problem* has an interesting and elegant solution. Example 6 below and the discussion in Section 7.5 will illustrate how such problems arise in practice.

The requirement that a vector $\mathbf{x}$ in $\mathbb{R}^n$ be a unit vector can be stated in several equivalent ways:

$$\|\mathbf{x}\| = 1, \qquad \|\mathbf{x}\|^2 = 1, \qquad \mathbf{x}^T\mathbf{x} = 1$$

and

$$x_1^2 + x_2^2 + \cdots + x_n^2 = 1 \tag{1}$$

The expanded version (1) of $\mathbf{x}^T\mathbf{x} = 1$ is commonly used in applications.

When a quadratic form $Q$ has no cross-product terms, it is easy to find the maximum and minimum of $Q(\mathbf{x})$ for $\mathbf{x}^T\mathbf{x} = 1$.

**EXAMPLE 1** Find the maximum and minimum values of $Q(\mathbf{x}) = 9x_1^2 + 4x_2^2 + 3x_3^2$ subject to the constraint $\mathbf{x}^T\mathbf{x} = 1$.

**SOLUTION** Since $x_2^2$ and $x_3^2$ are nonnegative, note that

$$4x_2^2 \le 9x_2^2 \qquad \text{and} \qquad 3x_3^2 \le 9x_3^2$$

and hence

$$\begin{aligned} Q(\mathbf{x}) &= 9x_1^2 + 4x_2^2 + 3x_3^2 \\ &\le 9x_1^2 + 9x_2^2 + 9x_3^2 \\ &= 9(x_1^2 + x_2^2 + x_3^2) \\ &= 9 \end{aligned}$$

whenever $x_1^2 + x_2^2 + x_3^2 = 1$. So the maximum value of $Q(\mathbf{x})$ cannot exceed 9 when $\mathbf{x}$ is a unit vector. Furthermore, $Q(\mathbf{x}) = 9$ when $\mathbf{x} = (1, 0, 0)$. Thus 9 is the maximum value of $Q(\mathbf{x})$ for $\mathbf{x}^T\mathbf{x} = 1$.

To find the minimum value of $Q(\mathbf{x})$, observe that

$$9x_1^2 \ge 3x_1^2, \qquad 4x_2^2 \ge 3x_2^2$$

and hence

$$Q(\mathbf{x}) \ge 3x_1^2 + 3x_2^2 + 3x_3^2 = 3(x_1^2 + x_2^2 + x_3^2) = 3$$

whenever $x_1^2 + x_2^2 + x_3^2 = 1$. Also, $Q(\mathbf{x}) = 3$ when $x_1 = 0$, $x_2 = 0$, and $x_3 = 1$. So 3 is the minimum value of $Q(\mathbf{x})$ when $\mathbf{x}^T\mathbf{x} = 1$. ■

It is easy to see in Example 1 that the matrix of the quadratic form $Q$ has eigenvalues 9, 4, and 3 and that the greatest and least eigenvalues equal, respectively, the (constrained) maximum and minimum of $Q(\mathbf{x})$. The same holds true for any quadratic form, as we shall see.

**EXAMPLE 2** Let $A = \begin{bmatrix} 3 & 0 \\ 0 & 7 \end{bmatrix}$, and let $Q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ for $\mathbf{x}$ in $\mathbb{R}^2$. Figure 1 displays the graph of $Q$. Figure 2 shows only the portion of the graph inside a cylinder; the intersection of the cylinder with the surface is the set of points $(x_1, x_2, z)$ such that $z = Q(x_1, x_2)$ and $x_1^2 + x_2^2 = 1$. The "heights" of these points are the constrained values of $Q(\mathbf{x})$. Geometrically, the constrained optimization problem is to locate the highest and lowest points on the intersection curve.

The two highest points on the curve are 7 units above the $x_1x_2$-plane, occurring where $x_1 = 0$ and $x_2 = \pm 1$. These points correspond to the eigenvalue 7 of $A$ and the eigenvectors $\mathbf{x} = (0, 1)$ and $-\mathbf{x} = (0, -1)$. Similarly, the two lowest points on the curve are 3 units above the $x_1x_2$-plane. They correspond to the eigenvalue 3 and the eigenvectors $(1, 0)$ and $(-1, 0)$. ■

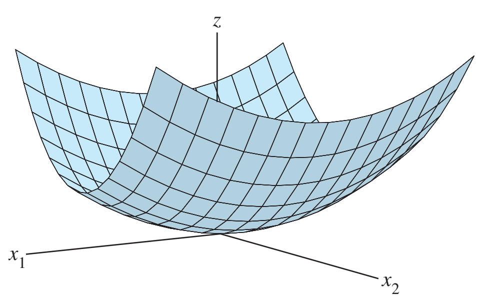



**FIGURE 1** $z = 3x_1^2 + 7x_2^2$.

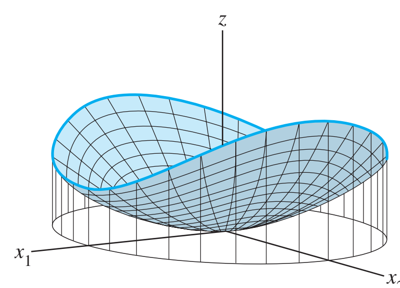



**FIGURE 2** The intersection of $z = 3x_1^2 + 7x_2^2$ and the cylinder $x_1^2 + x_2^2 = 1$.

Every point on the intersection curve in Fig. 2 has a $z$-coordinate between 3 and 7, and for any number $t$ between 3 and 7, there is a unit vector $\mathbf{x}$ such that $Q(\mathbf{x}) = t$. In other words, the set of all possible values of $\mathbf{x}^T A\mathbf{x}$, for $\|\mathbf{x}\| = 1$, is the closed interval $3 \le t \le 7$.

It can be shown that for any symmetric matrix $A$, the set of all possible values of $\mathbf{x}^T A\mathbf{x}$, for $\|\mathbf{x}\| = 1$, is a closed interval on the real axis. (See Exercise 13.) Denote the left and right endpoints of this interval by $m$ and $M$, respectively. That is, let

$$m = \min\{\mathbf{x}^T A\mathbf{x} : \|\mathbf{x}\| = 1\}, \quad M = \max\{\mathbf{x}^T A\mathbf{x} : \|\mathbf{x}\| = 1\} \tag{2}$$

Exercise 12 asks you to prove that if $\lambda$ is an eigenvalue of $A$, then $m \le \lambda \le M$. The next theorem says that $m$ and $M$ are themselves eigenvalues of $A$, just as in Example 2.[1]

**THEOREM 6**

Let $A$ be a symmetric matrix, and define $m$ and $M$ as in (2). Then $M$ is the greatest eigenvalue $\lambda_1$ of $A$ and $m$ is the least eigenvalue of $A$. The value of $\mathbf{x}^T A\mathbf{x}$ is $M$ when $\mathbf{x}$ is a unit eigenvector $\mathbf{u}_1$ corresponding to $M$. The value of $\mathbf{x}^T A\mathbf{x}$ is $m$ when $\mathbf{x}$ is a unit eigenvector corresponding to $m$.

[1] The use of *minimum* and *maximum* in (2), and *least* and *greatest* in the theorem, refers to the natural ordering of the real numbers, not to magnitudes.

**PROOF** Orthogonally diagonalize $A$ as $PDP^{-1}$. We know that

$$\mathbf{x}^T A\mathbf{x} = \mathbf{y}^T D\mathbf{y} \quad \text{when } \mathbf{x} = P\mathbf{y} \tag{3}$$

Also,

$$\|\mathbf{x}\| = \|P\mathbf{y}\| = \|\mathbf{y}\| \quad \text{for all } \mathbf{y}$$

because $P^TP = I$ and $\|P\mathbf{y}\|^2 = (P\mathbf{y})^T(P\mathbf{y}) = \mathbf{y}^TP^TP\mathbf{y} = \mathbf{y}^T\mathbf{y} = \|\mathbf{y}\|^2$. In particular, $\|\mathbf{y}\| = 1$ if and only if $\|\mathbf{x}\| = 1$. Thus $\mathbf{x}^TA\mathbf{x}$ and $\mathbf{y}^TD\mathbf{y}$ assume the same set of values as $\mathbf{x}$ and $\mathbf{y}$ range over the set of all unit vectors.

To simplify notation, suppose that $A$ is a $3 \times 3$ matrix with eigenvalues $a \geq b \geq c$. Arrange the (eigenvector) columns of $P$ so that $P = [\,\mathbf{u}_1 \quad \mathbf{u}_2 \quad \mathbf{u}_3\,]$ and

$$D = \begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix}$$

Given any unit vector $\mathbf{y}$ in $\mathbb{R}^3$ with coordinates $y_1, y_2, y_3$, observe that

$$\begin{aligned} ay_1^2 &= ay_1^2 \\ by_2^2 &\leq ay_2^2 \\ cy_3^2 &\leq ay_3^2 \end{aligned}$$

and obtain these inequalities:

$$\begin{aligned} \mathbf{y}^TD\mathbf{y} &= ay_1^2 + by_2^2 + cy_3^2 \\ &\leq ay_1^2 + ay_2^2 + ay_3^2 \\ &= a(y_1^2 + y_2^2 + y_3^2) \\ &= a\|\mathbf{y}\|^2 = a \end{aligned}$$

Thus $M \leq a$, by definition of $M$. However, $\mathbf{y}^TD\mathbf{y} = a$ when $\mathbf{y} = \mathbf{e}_1 = (1, 0, 0)$, so in fact $M = a$. By (3), the $\mathbf{x}$ that corresponds to $\mathbf{y} = \mathbf{e}_1$ is the eigenvector $\mathbf{u}_1$ of $A$, because

$$\mathbf{x} = P\mathbf{e}_1 = \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \mathbf{u}_3 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \mathbf{u}_1$$

Thus $M = a = \mathbf{e}_1^TD\mathbf{e}_1 = \mathbf{u}_1^TA\mathbf{u}_1$, which proves the statement about $M$. A similar argument shows that $m$ is the least eigenvalue, $c$, and this value of $\mathbf{x}^TA\mathbf{x}$ is attained when $\mathbf{x} = P\mathbf{e}_3 = \mathbf{u}_3$. ■

**EXAMPLE 3** Let $A = \begin{bmatrix} 3 & 2 & 1 \\ 2 & 3 & 1 \\ 1 & 1 & 4 \end{bmatrix}$. Find the maximum value of the quadratic form $\mathbf{x}^TA\mathbf{x}$ subject to the constraint $\mathbf{x}^T\mathbf{x} = 1$, and find a unit vector at which this maximum value is attained.

**SOLUTION** By Theorem 6, the desired maximum value is the greatest eigenvalue of $A$. The characteristic equation turns out to be

$$0 = -\lambda^3 + 10\lambda^2 - 27\lambda + 18 = -(\lambda - 6)(\lambda - 3)(\lambda - 1)$$

The greatest eigenvalue is 6.

The constrained maximum of $\mathbf{x}^TA\mathbf{x}$ is attained when $\mathbf{x}$ is a unit eigenvector for $\lambda = 6$. Solve $(A - 6I)\mathbf{x} = 0$ and find an eigenvector $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$. Set $\mathbf{u}_1 = \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}$. ■

In Theorem 7 and in later applications, the values of $\mathbf{x}^TA\mathbf{x}$ are computed with additional constraints on the unit vector $\mathbf{x}$.

**THEOREM 7**

Let $A$, $\lambda_1$, and $\mathbf{u}_1$ be as in Theorem 6. Then the maximum value of $\mathbf{x}^TA\mathbf{x}$ subject to the constraints

$$\mathbf{x}^T\mathbf{x} = 1, \quad \mathbf{x}^T\mathbf{u}_1 = 0$$

is the second greatest eigenvalue, $\lambda_2$, and this maximum is attained when $\mathbf{x}$ is an eigenvector $\mathbf{u}_2$ corresponding to $\lambda_2$.

Theorem 7 can be proved by an argument similar to the one above in which the theorem is reduced to the case where the matrix of the quadratic form is diagonal. The next example gives an idea of the proof for the case of a diagonal matrix.

**EXAMPLE 4** Find the maximum value of $9x_1^2 + 4x_2^2 + 3x_3^2$ subject to the constraints $\mathbf{x}^T\mathbf{x} = 1$ and $\mathbf{x}^T\mathbf{u}_1 = 0$, where $\mathbf{u}_1 = (1, 0, 0)$. Note that $\mathbf{u}_1$ is a unit eigenvector corresponding to the greatest eigenvalue $\lambda = 9$ of the matrix of the quadratic form.

**SOLUTION** If the coordinates of $\mathbf{x}$ are $x_1, x_2, x_3$, then the constraint $\mathbf{x}^T\mathbf{u}_1 = 0$ means simply that $x_1 = 0$. For such a unit vector, $x_2^2 + x_3^2 = 1$, and

$$\begin{aligned} 9x_1^2 + 4x_2^2 + 3x_3^2 &= 4x_2^2 + 3x_3^2 \\ &\le 4x_2^2 + 4x_3^2 \\ &= 4(x_2^2 + x_3^2) \\ &= 4 \end{aligned}$$

Thus the constrained maximum of the quadratic form does not exceed 4. And this value is attained for $\mathbf{x} = (0, 1, 0)$, which is an eigenvector for the second greatest eigenvalue of the matrix of the quadratic form. ■

**EXAMPLE 5** Let $A$ be the matrix in Example 3 and let $\mathbf{u}_1$ be a unit eigenvector corresponding to the greatest eigenvalue of $A$. Find the maximum value of $\mathbf{x}^TA\mathbf{x}$ subject to the conditions

$$\mathbf{x}^T\mathbf{x} = 1, \quad \mathbf{x}^T\mathbf{u}_1 = 0 \tag{4}$$

**SOLUTION** From Example 3, the second greatest eigenvalue of $A$ is $\lambda = 3$. Solve $(A - 3I)\mathbf{x} = \mathbf{0}$ to find an eigenvector, and normalize it to obtain

$$\mathbf{u}_2 = \begin{bmatrix} 1/\sqrt{6} \\ 1/\sqrt{6} \\ -2/\sqrt{6} \end{bmatrix}$$

The vector $\mathbf{u}_2$ is automatically orthogonal to $\mathbf{u}_1$ because the vectors correspond to different eigenvalues. Thus the maximum of $\mathbf{x}^TA\mathbf{x}$ subject to the constraints in (4) is 3, attained when $\mathbf{x} = \mathbf{u}_2$. ■

The next theorem generalizes Theorem 7 and, together with Theorem 6, gives a useful characterization of *all* the eigenvalues of $A$. The proof is omitted.

**THEOREM 8**

Let $A$ be a symmetric $n \times n$ matrix with an orthogonal diagonalization $A = PDP^{-1}$, where the entries on the diagonal of $D$ are arranged so that $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ and where the columns of $P$ are corresponding unit eigenvectors $\mathbf{u}_1, \ldots, \mathbf{u}_n$. Then for $k = 2, \ldots, n$, the maximum value of $\mathbf{x}^T A\mathbf{x}$ subject to the constraints

$$\mathbf{x}^T\mathbf{x} = 1, \quad \mathbf{x}^T\mathbf{u}_1 = 0, \quad \ldots, \quad \mathbf{x}^T\mathbf{u}_{k-1} = 0$$

is the eigenvalue $\lambda_k$, and this maximum is attained at $\mathbf{x} = \mathbf{u}_k$.

Theorem 8 will be helpful in Sections 7.4 and 7.5. The following application requires only Theorem 6.

**EXAMPLE 6** During the next year, a county government is planning to repair $x$ hundred miles of public roads and bridges and to improve $y$ hundred acres of parks and recreation areas. The county must decide how to allocate its resources (funds, equipment, labor, etc.) between these two projects. If it is more cost-effective to work simultaneously on both projects rather than on only one, then $x$ and $y$ might satisfy a *constraint* such as

$$4x^2 + 9y^2 \leq 36$$

See Fig. 3. Each point $(x, y)$ in the shaded *feasible set* represents a possible public works schedule for the year. The points on the constraint curve, $4x^2 + 9y^2 = 36$, use the maximum amounts of resources available.

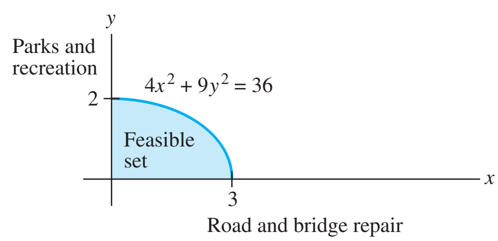



**FIGURE 3** Public works schedules.

In choosing its public works schedule, the county wants to consider the opinions of the county residents. To measure the value, or *utility*, that the residents would assign to the various work schedules $(x, y)$, economists sometimes use a function such as

$$q(x, y) = xy$$

The set of points $(x, y)$ at which $q(x, y)$ is a constant is called an *indifference curve*. Three such curves are shown in Fig. 4. Points along an indifference curve correspond to alternatives that county residents as a group would find equally valuable.[2] Find the public works schedule that maximizes the utility function $q$.

**SOLUTION** The constraint equation $4x^2 + 9y^2 = 36$ does not describe a set of unit vectors, but a change of variable can fix that problem. Rewrite the constraint in the form

$$\left(\frac{x}{3}\right)^2 + \left(\frac{y}{2}\right)^2 = 1$$

[2] Indifference curves are discussed in Michael D. Intriligator, Ronald G. Bodkin, and Cheng Hsiao, *Econometric Models, Techniques, and Applications* (Upper Saddle River, NJ: Prentice-Hall, 1996).

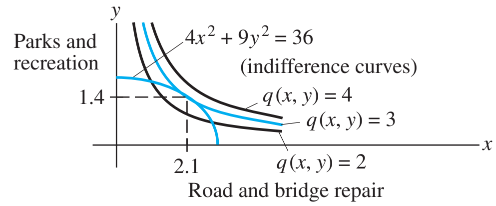



**FIGURE 4** The optimum public works schedule is (2.1, 1.4).

and define

$$x_1 = \frac{x}{3}, \quad x_2 = \frac{y}{2}, \qquad \text{that is,} \quad x = 3x_1 \quad \text{and} \quad y = 2x_2$$

Then the constraint equation becomes

$$x_1^2 + x_2^2 = 1$$

and the utility function becomes $q(3x_1, 2x_2) = (3x_1)(2x_2) = 6x_1x_2$. Let $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$. Then the problem is to maximize $Q(\mathbf{x}) = 6x_1x_2$ subject to $\mathbf{x}^T\mathbf{x} = 1$. Note that $Q(\mathbf{x}) = \mathbf{x}^TA\mathbf{x}$, where

$$A = \begin{bmatrix} 0 & 3 \\ 3 & 0 \end{bmatrix}$$

The eigenvalues of $A$ are $\pm 3$, with eigenvectors $\begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$ for $\lambda = 3$ and $\begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$ for $\lambda = -3$. Thus the maximum value of $Q(\mathbf{x}) = q(x_1, x_2)$ is 3, attained when $x_1 = 1/\sqrt{2}$ and $x_2 = 1/\sqrt{2}$.

In terms of the original variables, the optimum public works schedule is $x = 3x_1 = 3/\sqrt{2} \approx 2.1$ hundred miles of roads and bridges and $y = 2x_2 = \sqrt{2} \approx 1.4$ hundred acres of parks and recreational areas. The optimum public works schedule is the point where the constraint curve and the indifference curve $q(x, y) = 3$ just meet. Points $(x, y)$ with a higher utility lie on indifference curves that do not touch the constraint curve. See Fig. 4. ■

## PRACTICE PROBLEMS

1. Let $Q(\mathbf{x}) = 3x_1^2 + 3x_2^2 + 2x_1x_2$. Find a change of variable that transforms $Q$ into a quadratic form with no cross-product term, and give the new quadratic form.
2. With $Q$ as in Problem 1, find the maximum value of $Q(\mathbf{x})$ subject to the constraint $\mathbf{x}^T\mathbf{x} = 1$, and find a unit vector at which the maximum is attained.

# 7.3 EXERCISES

In Exercises 1 and 2, find the change of variable $\mathbf{x} = P\mathbf{y}$ that transforms the quadratic form $\mathbf{x}^TA\mathbf{x}$ into $\mathbf{y}^TD\mathbf{y}$ as shown.

1. $5x_1^2 + 6x_2^2 + 7x_3^2 + 4x_1x_2 - 4x_2x_3 = 9y_1^2 + 6y_2^2 + 3y_3^2$
2. $3x_1^2 + 2x_2^2 + 2x_3^2 + 2x_1x_2 + 2x_1x_3 + 4x_2x_3 = 5y_1^2 + 2y_2^2$
[*Hint:* $\mathbf{x}$ and $\mathbf{y}$ must have the same number of coordinates, so the quadratic form shown here must have a coefficient of zero for $y_3^2$.]

In Exercises 3–6, find (a) the maximum value of $Q(\mathbf{x})$ subject to the constraint $\mathbf{x}^T\mathbf{x} = 1$, (b) a unit vector $\mathbf{u}$ where this maximum is attained, and (c) the maximum of $Q(\mathbf{x})$ subject to the constraints $\mathbf{x}^T\mathbf{x} = 1$ and $\mathbf{x}^T\mathbf{u} = 0$.

3. $Q(\mathbf{x}) = 5x_1^2 + 6x_2^2 + 7x_3^2 + 4x_1x_2 - 4x_2x_3$
(See Exercise 1.)

**4.** $Q(\mathbf{x}) = 3x_1^2 + 2x_2^2 + 2x_3^2 + 2x_1x_2 + 2x_1x_3 + 4x_2x_3$ (See Exercise 2.)

**5.** $Q(\mathbf{x}) = 5x_1^2 + 5x_2^2 - 4x_1x_2$

**6.** $Q(\mathbf{x}) = 7x_1^2 + 3x_2^2 + 3x_1x_2$

**7.** Let $Q(\mathbf{x}) = -2x_1^2 - x_2^2 + 4x_1x_2 + 4x_2x_3$. Find a unit vector $\mathbf{x}$ in $\mathbb{R}^3$ at which $Q(\mathbf{x})$ is maximized, subject to $\mathbf{x}^T\mathbf{x} = 1$. [*Hint:* The eigenvalues of the matrix of the quadratic form $Q$ are 2, $-1$, and $-4$.]

**8.** Let $Q(\mathbf{x}) = 7x_1^2 + x_2^2 + 7x_3^2 - 8x_1x_2 - 4x_1x_3 - 8x_2x_3$. Find a unit vector $\mathbf{x}$ in $\mathbb{R}^3$ at which $Q(\mathbf{x})$ is maximized, subject to $\mathbf{x}^T\mathbf{x} = 1$. [*Hint:* The eigenvalues of the matrix of the quadratic form $Q$ are 9 and $-3$.]

**9.** Find the maximum value of $Q(\mathbf{x}) = 7x_1^2 + 3x_2^2 - 2x_1x_2$, subject to the constraint $x_1^2 + x_2^2 = 1$. (Do not go on to find a vector where the maximum is attained.)

**10.** Find the maximum value of $Q(\mathbf{x}) = -3x_1^2 + 5x_2^2 - 2x_1x_2$, subject to the constraint $x_1^2 + x_2^2 = 1$. (Do not go on to find a vector where the maximum is attained.)

**11.** Suppose $\mathbf{x}$ is a unit eigenvector of a matrix $A$ corresponding to an eigenvalue 3. What is the value of $\mathbf{x}^TA\mathbf{x}$?

**12.** Let $\lambda$ be any eigenvalue of a symmetric matrix $A$. Justify the statement made in this section that $m \le \lambda \le M$, where $m$ and $M$ are defined as in (2). [*Hint:* Find an $\mathbf{x}$ such that $\lambda = \mathbf{x}^TA\mathbf{x}$.]

**13.** Let $A$ be an $n \times n$ symmetric matrix, let $M$ and $m$ denote the maximum and minimum values of the quadratic form $\mathbf{x}^TA\mathbf{x}$, and denote corresponding unit eigenvectors by $\mathbf{u}_1$ and $\mathbf{u}_n$. The following calculations show that given any number $t$ between $M$ and $m$, there is a unit vector $\mathbf{x}$ such that $t = \mathbf{x}^TA\mathbf{x}$. Verify that $t = (1-\alpha)m + \alpha M$ for some number $\alpha$ between 0 and 1. Then let $\mathbf{x} = \sqrt{1-\alpha}\mathbf{u}_n + \sqrt{\alpha}\mathbf{u}_1$, and show that $\mathbf{x}^T\mathbf{x} = 1$ and $\mathbf{x}^TA\mathbf{x} = t$.

[**M**] In Exercises 14–17, follow the instructions given for Exercises 3–6.

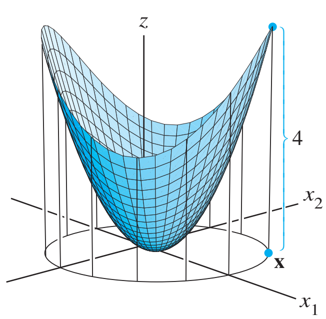
**14.** $x_1x_2 + 3x_1x_3 + 30x_1x_4 + 30x_2x_3 + 3x_2x_4 + x_3x_4$

**15.** $3x_1x_2 + 5x_1x_3 + 7x_1x_4 + 7x_2x_3 + 5x_2x_4 + 3x_3x_4$

**16.** $4x_1^2 - 6x_1x_2 - 10x_1x_3 - 10x_1x_4 - 6x_2x_3 - 6x_2x_4 - 2x_3x_4$

**17.** $-6x_1^2 - 10x_2^2 - 13x_3^2 - 13x_4^2 - 4x_1x_2 - 4x_1x_3 - 4x_1x_4 + 6x_3x_4$

## SOLUTIONS TO PRACTICE PROBLEMS




The maximum value of $Q(\mathbf{x})$ subject to $\mathbf{x}^T\mathbf{x} = 1$ is 4.

**1.** The matrix of the quadratic form is $A = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$. It is easy to find the eigenvalues, 4 and 2, and corresponding unit eigenvectors, $\begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$ and $\begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$. So the desired change of variable is $\mathbf{x} = P\mathbf{y}$, where $P = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$. (A common error here is to forget to normalize the eigenvectors.) The new quadratic form is $\mathbf{y}^TD\mathbf{y} = 4y_1^2 + 2y_2^2$.

**2.** The maximum of $Q(\mathbf{x})$ for $\mathbf{x}$ a unit vector is 4, and the maximum is attained at the unit eigenvector $\begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$. [A common incorrect answer is $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$. This vector maximizes the quadratic form $\mathbf{y}^TD\mathbf{y}$ instead of $Q(\mathbf{x})$.]

# 7.4 THE SINGULAR VALUE DECOMPOSITION

The diagonalization theorems in Sections 5.3 and 7.1 play a part in many interesting applications. Unfortunately, as we know, not all matrices can be factored as $A = PDP^{-1}$ with $D$ diagonal. However, a factorization $A = QDP^{-1}$ *is* possible for any $m \times n$ matrix $A$! A special factorization of this type, called the *singular value decomposition*, is one of the most useful matrix factorizations in applied linear algebra.

The singular value decomposition is based on the following property of the ordinary diagonalization that can be imitated for rectangular matrices: The absolute values of the eigenvalues of a symmetric matrix $A$ measure the amounts that $A$ stretches or shrinks